



REC'D 26 SEP 2000
WIPO PCT

BREVET D'INVENTION

10/049372

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

#2
(Handwritten signature)

DOCUMENT DE PRIORITÉ

COPIE OFFICIELLE

PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA
RÈGLE 17.1.a) OU b)

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 04 SEP. 2000

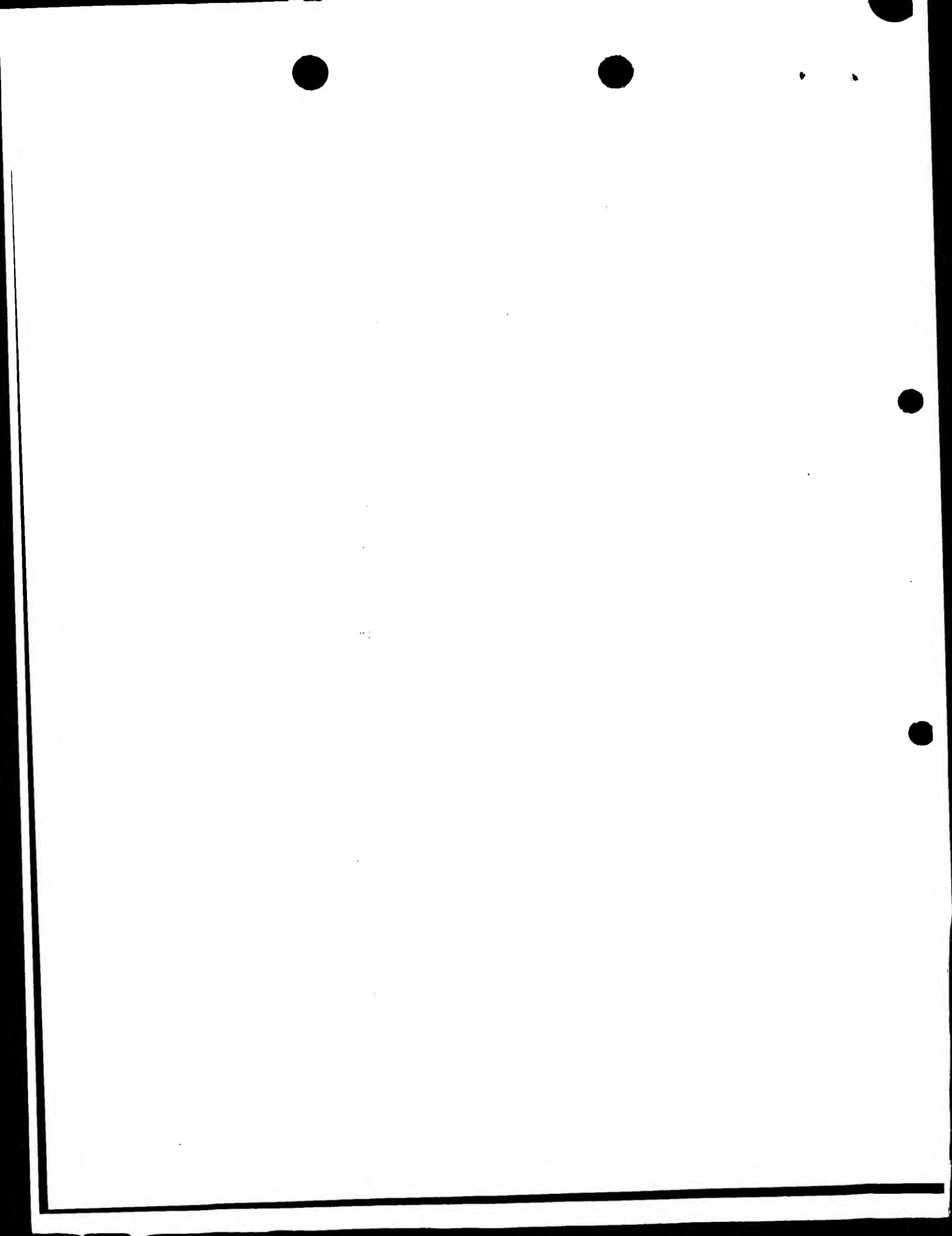
Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

M. Heuc

Martine PLANCHE

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint Petersbourg
75800 PARIS Cedex 08
Téléphone : 01 53 04 53 04
Télécopie : 01 42 93 59 30



26 bis, rue de Saint-Pétersbourg

75800 Paris Cedex 08

Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

REQUÊTE EN DÉLIVRANCE

Confirmation d'un dépôt par télécopie

Cet imprimé est à remplir à l'encre noire en lettres capitales

Réserve à l'INPI

DATE DE REMISE DES PIÈCES

12 AOUT 1999

N° D'ENREGISTREMENT NATIONAL

9910439

DÉPARTEMENT DE DÉPÔT

75 INPI PARIS

DATE DE DÉPÔT

12 AOUT 1999

1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE
À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE

CABINET REGIMBEAU
26, Avenue Kléber
75116 PARIS

n° du pouvoir permanent références du correspondant téléphone
237913 D18275 PM **01 45 00 92 02**

date

2 DEMANDE Nature du titre de propriété industrielle

- brevet d'invention demande divisionnaire
 certificat d'utilité transformation d'une demande
de brevet européen
 brevet d'invention



Établissement du rapport de recherche

différé immédiat

Le demandeur, personne physique, requiert le paiement échelonné de la redevance

oui non

Titre de l'invention (200 caractères maximum)

"Odorant-Binding" Protéines humaines fixant des ligands hydrophobes : polypeptides et polynucléotides codant lesdits polypeptides et leurs applications.

3 DEMANDEUR (S) n° SIREN

code APE-NAF

Nom et prénoms (souligner le nom patronymique) ou dénomination

Forme juridique

1. UNIVERSITÉ D'AUVERGNE
2. PITIOT Gilles

Nationalité (s) **Française, Française**

Adresse (s) complète (s)

Pays

1. 49, boulevard François Mitterrand, 63000 CLERMONT-FERRAND,
2. 151, rue du Chevaleret 75013 PARIS,

FR

FR

En cas d'insuffisance de place, poursuivre sur papier libre

4 INVENTEUR (S) Les inventeurs sont les demandeurs

oui

non

Si la réponse est non, fournir une désignation séparée

5 RÉDUCTION DU TAUX DES REDEVANCES

requise pour la 1ère fois

requise antérieurement au dépôt : joindre copie de la décision d'admission

6 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE

pays d'origine

numéro

date de dépôt

nature de la demande

7 DIVISIONS antérieures à la présente demande n°

date

n°

date

8 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE

(nom et qualité du signataire)

SIGNATURE DU PRÉPOSÉ À LA RÉCEPTION

SIGNATURE APRÈS ENREGISTREMENT DE LA DEMANDE À L'INPI

92-1001

DEPARTEMENT DES BREVETS

26bis, rue de Saint-Pétersbourg
75800 Paris Cedex 08
Tél. : 01 53 04 53 04 - Télécopie : 01 42 93 59 30

DÉSIGNATION DE L'INVENTEUR
(si le demandeur n'est pas l'inventeur ou l'unique inventeur)

N° D'ENREGISTREMENT NATIONAL

99 10439

TITRE DE L'INVENTION : "Odorant-Binding" Protéines humaines fixant des ligands hydrophobes : polypeptides et polynucléotides codant lesdits polypeptides et leurs applications.

LE(S) SOUSSIGNÉ(S)

UNIVERSITE D'AUVERGNE
PITIOT Gilles
49, boulevard François Mitterrand, 63000 CLERMONT-FERRAND,
151, rue du Chevaleret 75013 PARIS, FR

DÉSIGNE(NT) EN TANT QU'INVENTEUR(S) (indiquer nom, prénoms, adresse et souligner le nom patronymique) :

PITIOT Gilles
151, rue Chevaleret
75013 PARIS, FR

LACAZETTE Eric
66, Bd La Fayette
Appt 211
63000 CLERMONT-FERRAND
FR

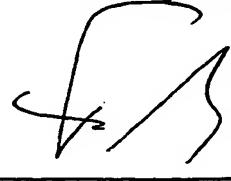
GACHON Françoise
11, rue des Paillards
63540 ROMAGNAT, FR

NOTA : A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient (société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.

Date et signature (s) du (des) demandeur (s) ou du mandataire

23 septembre 1999

CABINET REGIMBEAU


g2-1001

DOCUMENT COMPORTANT DES MODIFICATIONS

PAGE(S) DE LA DESCRIPTION OU DES REVENDICATIONS OU PLANCHE(S) DE DESSIN			R.M.*	DATE DE LA CORRESPONDANCE	TAMPON DATEUR DU CORRECTEUR
Modifiée(s)	Supprimée(s)	Ajoutée(s)			
p 94-101			α	7/2/2000	VD - 3/2/2000
p 12-A4				7/2/2000	VD - 3/2/2000

Un changement apporté à la rédaction des revendications d'origine, sauf si celui-ci découle des dispositions de l'article R.612-36 du code de la Propriété Intellectuelle, est signalé par la mention «R.M.» (revendications modifiées).

La présente invention concerne la mise en évidence de nouvelles protéines humaines fixant les odeurs, ci-après dénommées « OBP » (Odorant Binding Proteins) ainsi que leurs applications, tant au niveau thérapeutique que non thérapeutique.

5 La présente invention repose sur l'identification d'une famille de gènes de lipocalines composée de trois gènes et de deux pseudogènes sur le chromosome humain 9q34 ; les trois gènes correspondent au gène LCN1 déjà décrit et à deux nouveaux gènes faisant l'objet de la présente invention et dénommés hOBPIIa et hOBPIIb. De plus, 10 l'invention repose sur l'attribution de nouvelles fonctions à des lipocalines humaines déjà connues par la mise en évidence de nouveaux territoires d'expression.

Bien qu'un certain nombre d'OBP ait déjà été mis en évidence (Pelosi *et al.*, 1996), telle l'OBP de rat par exemple (voir notamment le 15 brevet EP-0 335 654), les OBP selon la présente invention sont des OBP humaines qui présentent un très grand nombre d'avantages, comme cela ressortira de la suite du texte, par rapport aux protéines murines.

Ces protéines OBP de la famille des lipocalines ont, pour 20 certaines d'entre elles, été mentionnées indirectement dans le brevet WO 99 07740, mais leur fonction d'OBP n'a jamais été décrite jusqu'à présent ; il en est ainsi également des protéines LCN1, rétinol-binding-protein (RBP) et Apolipoprotéine D (ApoD), comme cela sera explicité plus complètement ci-après.

25 Historiquement, la famille des lipocalines (Pervaiz et Brew, 1987) a été définie à partir de la protéine humaine fixant le rétinol (RBP) et à partir de 3 autres protéines : la β -lactoglobuline de boeuf, l' α 2 μ -globuline de rat et l' α 1-microglobuline humaine. A partir de ces protéines de référence et en utilisant des homologies de séquence, la 30 famille des lipocalines s'est enrichie pour incorporer maintenant un grand nombre de protéines, plus d'une centaine, aussi bien chez les

eucaryotes que chez les procaryotes (Flower *et al.*, 1995, 1996). Cette famille consiste en de petites protéines (160-190 acides aminés) contenant une poche hydrophobe et qui sont généralement sécrétées (Bocskei *et al.*, 1992 ; Senoo *et al.*, 1990, Zeng *et al.*, 1996 ; Miller, 1998), bien que dans certains cas il s'agisse de protéines qui restent associées à la membrane (Nagata *et al.*, 1991). Chez les vertébrés, les identités de séquence entre les différentes lipocalines tournent autour de 20%, toutefois, les identités de séquences sont plus importantes pour les protéines orthologues, de même que pour les récents gènes paralogues décrits (Igarashi *et al.*, 1992 ; Dewald *et al.*, 1992).

La présente invention concerne les polypeptides isolés OBPII codés par les deux nouveaux gènes humains OBPIIa et OBPIIb isolés localisés au locus 9q34 ; l'invention concerne également les séquences polynucléotidiques correspondantes, les ARNm correspondants, ainsi que les séquences régulatrices promotrices qui déterminent le profil d'expression dans les différents tissus et notamment dans les tissus sécréteurs. Ces deux gènes codent pour au moins sept polypeptides différents étant donné l'existence d'épissage alternatif des transcrits. Le gène OBPIIa code pour au moins 4 polypeptides différents dénommés OBPII_{aα}, OBPII_{aβ}, OBPII_{aγ}, OBPII_{aδ} et le gène OBPIIb code pour trois polypeptides différents dénommés OBPII_{bα}, OBPII_{bβ}, OBPII_{bδ}. Même si un gène est principalement exprimé, toutes les formes de messagers sont retrouvées dans la structure nasale.

La présente invention a donc pour objet un polypeptide isolé comprenant une séquence en amino-acides ayant au moins 90% d'identité avec les séquences en amino-acides SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14 et LCN1, ApoD et RBP.

Il doit être compris que l'invention concerne les polypeptides obtenus par purification à partir de sources naturelles ou bien obtenus par les techniques recombinantes, comme cela sera décrit dans ce qui

va suivre; l'invention concerne également les polypeptides obtenus par synthèse chimique qui peuvent alors comporter des acides aminés non naturels. Dans la présente description, on utilisera le terme polypeptide pour désigner également une protéine ou un peptide.

5 Par pourcentage d'identité entre les séquences, on entend désigner le pourcentage d'acides aminés identiques entre les séquences obtenu avec le meilleur alignement de séquences possibles. Ce pourcentage étant purement statistique et les différences entre les polypeptides étant réparties au hasard et sur toute la longueur.

10 La présente invention concerne également un polypeptide isolé caractérisé en ce qu'il comprend un polypeptide choisi parmi :

- a) un polypeptide de séquence SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14;
- 15 b) un polypeptide variant de polypeptide de séquences d'acides aminés défini en a) ;
- c) un polypeptide homologue au polypeptide défini en a) ou b) et comportant au moins 90 % d'identité avec ledit polypeptide de a) ;
- d) un fragment d'au moins 15 acides aminés consécutifs d'un polypeptide défini en a), b) ou c) ;
- 20 e) un fragment biologiquement actif d'un polypeptide défini en a), b) ou c).

On entendra par polypeptide variant l'ensemble des polypeptides mutés pouvant exister naturellement, en particulier chez l'être humain, et qui correspondent notamment à des troncatures, 25 substitutions, délétions et/ou additions de résidus d'amino-acides. Les polypeptides variants selon l'invention conservent au moins un domaine de fixation à un ligand hydrophobe.

Par polypeptide homologue, on entendra désigner les polypeptides présentant, par rapport aux polypeptides de séquence 30 SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 et SEQ ID N°14, certaines modifications comme en particulier

une délétion, addition ou substitution d'au moins un acide aminé, une troncature, un allongement et/ou une fusion chimérique. Parmi les polypeptides homologues, on préfère ceux dont la séquence d'acides aminés présente au moins 90 % d'identité, de préférence 95 %, de manière préférée 97 %, et de manière encore préférée 99 % d'identité avec les séquences d'acides aminés des polypeptides selon l'invention.

Dans le cas d'une substitution, un ou plusieurs acides aminés consécutifs ou non consécutifs, sont remplacés par des « acides aminés équivalents ». L'expression « acide aminé équivalent » vise ici à désigner tout acide aminé susceptible d'être substitué à l'un des acides aminés de la structure de base sans cependant modifier les caractéristiques ou propriétés fonctionnelles essentielles des polypeptides correspondants, leurs activités biologiques, telles par exemple l'induction *in vivo* d'anticorps capables de reconnaître le polypeptide dont la séquence d'acides aminés est comprise dans la séquence d'acides aminés SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14, ou l'un de ses fragments. Ces acides aminés équivalents peuvent être déterminés soit en s'appuyant sur leur homologie de structure avec les acides aminés auxquels ils se substituent, soit sur les résultats des essais d'activité biologique croisée auxquels les différents polypeptides sont susceptibles de donner lieu. A titre d'exemple, on mentionnera les possibilités de substitutions susceptibles d'être effectuées sans qu'il en résulte une modification approfondie des activités biologiques des polypeptides modifiés correspondants, les remplacements, par exemple, de la leucine par la valine ou l'isoleucine, de l'acide aspartique par l'acide glutamique, de la glutamine par l'asparagine, de l'arginine par la lysine etc., les substitutions inverses étant naturellement envisageables dans les mêmes conditions. Ainsi, on peut envisager d'introduire certaines modifications comme en particulier une délétion, addition ou substitution d'au moins un acide aminé au niveau des hélices alpha de

la protéine sans détruire le calice formé par la structure composée des feuillets beta ; de la même manière, il est possible d'introduire des amino-acides équivalents qui permettent de conserver aux feuillets bêta le caractère hydrophobe. Il peut également être intéressant 5 d'introduire des modifications dans la séquence des polypeptides de l'invention pour générer des polypeptides homologues dépourvus de sites protéasiques.

Par fragment biologiquement actif, on entendra désigner en particulier un fragment de séquence d'acides aminés de polypeptide 10 selon l'invention présentant au moins une des caractéristiques ou propriétés fonctionnelles des polypeptides selon l'invention, notamment en ce que : (i) il est capable d'être reconnu par un anticorps spécifique d'un polypeptide selon l'invention ou bien par des anticorps produits par des patients au cours d'une réaction 15 immunitaire; (ii) il présente au moins l'un des domaines ou régions tels que définis ci-après ; (iii) il est capable de lier un ligand hydrophobe et notamment des molécules odorantes, de préférence les phéromones ; (iv) il est capable de lier spécifiquement un récepteur.

Par fragment de polypeptide, on entend désigner un polypeptide 20 comportant au minimum 15 acides aminés, de préférence 18 acides aminés, de manière préférée 25 et de manière encore préférée 50 acides aminés. Les fragments de polypeptide selon l'invention obtenus par clivage dudit polypeptide par une enzyme protéolytique, par un réactif chimique, ou encore en plaçant ledit polypeptide dans un 25 environnement très acide font également partie de l'invention. Lorsque l'on souhaite utiliser une séquence de 15, 18, 25 ou de 50 amino-acides, il s'agit bien entendu, de préférence, des parties correspondant à des épitopes fonctionnels des polypeptides précédents, lesquels pourront être intégrés dans un polypeptide ayant une structure plus 30 longue sous forme par exemple de protéine de fusion, ceci dépendra des applications que l'on souhaite de ces polypeptides.

Selon un mode préféré de réalisation, la présente invention concerne un polypeptide isolé sélectionné parmi un polypeptide correspondant à la séquence SEQ ID N°2 et dénommé OBPII_{αα}, à la séquence SEQ ID N°4 et dénommé OBPII_{αβ}, à la séquence SEQ ID N°6 et dénommé OBPII_{αγ}, à la séquence SEQ ID N°10 et dénommé OBPII_{βα}, à la séquence SEQ ID N°12 et dénommé OBPII_{ββ}.

Les polypeptides selon l'invention se caractérisent en ce qu'ils comportent de préférence le domaine Gly-Thr-Trp-Tyr.

L'invention concerne également le polynucléotide isolé caractérisé en ce qu'il code pour un polypeptide tel que décrit précédemment. Le polynucléotide selon l'invention est choisi parmi le polynucléotide de séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 ou SEQ ID N° 13.

L'invention porte également sur un polynucléotide isolé caractérisé en ce qu'il comprend un polynucléotide choisi parmi :

a) un polynucléotide de séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 ou SEQ ID N° 13 ou dont la séquence est celle de l'ARN correspondant à la séquence SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 ou SEQ ID N° 13 ;

b) un polynucléotide dont la séquence est complémentaire de la séquence d'un polynucléotide défini en a),

c) un polynucléotide dont la séquence comporte au moins 80% d'identité, de préférence 90 %, de manière préférée 95 %, et de manière encore préférée 97 % d'identité avec un polynucléotide défini en a) ou b),

d) un polynucléotide hybride dans des conditions de forte stringence avec une séquence de polynucléotide défini en a), b) ou c),

e) un fragment d'au moins 15 nucléotides consécutifs, de préférence 21 nucléotides consécutifs, et de manière préférée 30 nucléotides consécutifs d'un polynucléotide défini en a), b), c) ou d).

Dans la présente description, on entendra désigner par polynucléotide, oligonucléotide, séquence de polynucléotide, séquence nucléotidique ou acide nucléique un fragment d'ADN, aussi bien un ADN double brin, un ADN simple brin que des produits de transcription desdits ADN, et/ou un fragment d'ARN, lesdits fragments naturels isolés, ou de synthèse, comportant ou non des nucléotides non naturels, désignant un enchaînement précis de nucléotides, modifiés ou non, permettant de définir un fragment ou une région d'un acide nucléique.

Par polynucléotide de séquence complémentaire, on entend désigner tout ADN dont les nucléotides sont complémentaires de ceux de la SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11, SEQ ID N° 13 ou d'une partie de la SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11, SEQ ID N° 13 et dont l'orientation est inversée.

Par pourcentage d'identité au sens de la présente invention, on entend un pourcentage de bases identiques entre les polynucléotides obtenus après le meilleur alignement, ce pourcentage étant purement statistique et les différences entre les deux polynucléotides étant réparties au hasard et sur toute leur longueur.

Une hybridation dans des conditions de forte stringence signifie au sens de la présente invention, que les conditions de température et de force ionique sont choisies de telle manière qu'elles permettent le maintien de l'hybridation entre deux fragments d'ADN complémentaires. A titre illustratif, des conditions de forte stringence de l'étape d'hybridation aux fins de définir les fragments polynucléotidiques décrits ci-dessus, sont avantageusement les suivantes :

L'hybridation ADN-ADN ou ADN-ARN est réalisée en deux étapes: (1) préhybridation à 42°C pendant 3 heures en tampon phosphate (20 mM, pH 7,5) contenant 5 x SSC (1 x SSC correspond à

une solution 0,15 M NaCl + 0,015 M citrate de sodium), 50 % de formamide, 7 % de sodium dodécyl sulfate (SDS), 10 x Denhard's, 5 % de dextran sulfate et 100 µg/ml d'ADN de sperme de saumon ; (2) hybridation proprement dite pendant 20 heures à une température dépendant de la taille de la sonde (i.e. : 42°C, pour une sonde de taille > 100 nucléotides) suivie de 2 lavages de 20 minutes à 20°C en 2 x SSC + 2 % SDS, 1 lavage de 20 minutes à 20°C en 0,1 x SSC + 0,1 % SDS. Le dernier lavage est pratiqué en 0,1 x SSC + 0,1 % SDS pendant 30 minutes à 60°C pour une sonde de taille > 100 nucléotides. Les conditions d'hybridation de forte stringence décrites ci-dessus pour un polynucléotide de taille définie, seront adaptées par l'homme du métier pour des oligonucléotides de taille plus grande ou plus petite, selon l'enseignement de Sambrook *et al.*(1989).

Avantageusement, un fragment nucléotidique répondant à la définition précédente aura au moins 15 nucléotides consécutifs, de préférence au moins 21 nucléotides, et encore plus préférentiellement au moins 30 nucléotides consécutifs de la séquence dont il est issu.

Selon un mode de réalisation de l'invention, le polynucléotide selon l'invention se caractérise en ce qu'il est marqué directement ou indirectement par un composé radioactif ou un composé non radioactif. Le polynucléotide selon l'invention est utilisé en tant qu'amorce pour l'amplification ou la polymérisation de séquences nucléiques ; l'invention porte également sur l'utilisation d'un polynucléotide selon l'invention en tant que sonde pour la détection de séquences nucléiques. Selon l'invention, les fragments de polynucléotides pouvant être utilisés comme sonde ou comme amorce dans des procédés de détection, d'identification, de dosage ou d'amplification de séquence nucléique, présenteront une taille minimale de 9 bases, de préférence de 18 bases, et de manière plus préférée 36 bases. Enfin, l'invention porte sur l'utilisation d'un polynucléotide selon l'invention en tant que oligonucléotide sens ou

antisens pour contrôler l'expression du produit protéique correspondant en l'occurrence un polypeptide selon l'invention.

Les séquences de polynucléotides selon l'invention non marquées peuvent être utilisées directement comme sonde, amorce ou 5 oligonucléotide ; cependant les séquences utilisées sont généralement marquées pour obtenir des séquences utilisables pour de nombreuses applications. Le marquage des amorces, des sondes, des oligonucléotides selon l'invention est réalisé par des éléments radioactifs ou par des molécules non radioactives ; parmi les isotopes 10 radioactifs utilisés, on peut citer le ^{32}P , le ^{33}P , le ^{35}S , le ^3H ou le ^{125}I . Les entités non radioactives sont sélectionnées parmi les ligands tels la biotine, l'avidine, la streptavidine, la dioxygénine, les haptènes, les colorants, les agents luminescents tels que les agents 15 radioluminescents, chémiluminescents, bioluminescents, fluorescents, phosphorescents.

L'invention comprend également une méthode de détection et/ou de dosage d'acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes : (i) d'isolement de l'ADN à partir de l'échantillon biologique à 20 analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ; (ii) d'amplification spécifique de l'ADN codant pour le polypeptide selon l'invention à l'aide d'amorces ; (iii) d'analyse des produits d'amplification.

L'invention comprend en outre un nécessaire pour la détection 25 et/ou le dosage d'un acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) un couple d'amorces nucléiques selon l'invention, (ii) les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN, et éventuellement (iii) un composant permettant de vérifier la séquence 30 du fragment amplifié, plus particulièrement une sonde selon l'invention.

L'invention comprend aussi une méthode de détection et/ou de dosage d'acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes : (i) de mise en contact d'une sonde selon l'invention avec un échantillon biologique ; (ii) de détection et/ou de dosage de l'hybride formé entre ladite sonde et l'ADN de l'échantillon biologique.

L'invention comprend également un nécessaire pour la détection et/ou le dosage d'acide nucléique selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) une sonde selon l'invention, (ii) les réactifs nécessaires à la mise en oeuvre d'une réaction d'hybridation, et le cas échéant, (iii) un couple d'amorces selon l'invention, ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN.

L'invention concerne particulièrement les procédés selon l'invention et décrits ci-dessus, pour la détection et le diagnostic de cellules d'origine cancéreuse et principalement cancers du sein, de l'utérus, de l'ovaire, de la prostate et du poumon.

Les polynucléotides selon l'invention peuvent ainsi être utilisés comme amorce et/ou sonde dans des procédés mettant en oeuvre notamment la technique PCR (réaction en chaîne à la polymérase)(Erlich, 1989 ; Innis *et al.*, 1990, et Rolfs *et al.*, 1991). Cette technique nécessite le choix de paires d'amorces oligonucléotidiques encadrant le fragment qui doit être amplifié. On peut, par exemple, se référer à la technique décrite dans le brevet américain U.S. N° 4 683 202. Les fragments amplifiés peuvent être identifiés, par exemple après une électrophorèse en gel d'agarose ou de polyacrylamide, ou après une technique chromatographique comme la filtration sur gel ou la chromatographie échangeuse d'ions. La spécificité de l'amplification peut être contrôlée par hybridation moléculaire en utilisant comme sonde les séquences nucléotidiques de polynucléotides de l'invention, des plasmides contenant ces séquences

ou leurs produits d'amplification. Les fragments nucléotidiques amplifiés peuvent être utilisés comme réactifs dans des réactions d'hybridation afin de mettre en évidence la présence, dans un échantillon biologique, d'un acide nucléique cible de séquence complémentaire à celle desdits fragments nucléotidiques amplifiés.

5 L'invention vise également les fragments nucléotidiques susceptibles d'être obtenus par amplification à l'aide d'amorces selon l'invention.

D'autres techniques d'amplification de l'acide nucléique cible 10 peuvent être avantageusement employées comme alternative à la PCR (PCR-like) à l'aide de couple d'amorces de séquences nucléotidiques selon l'invention. Par PCR-like on entendra désigner toutes les méthodes mettant en oeuvre des reproductions directes ou indirectes 15 des séquences d'acides nucléiques, ou bien dans lesquelles les systèmes de marquage ont été amplifiés, ces techniques sont bien entendu connues, en général il s'agit de l'amplification de l'ADN par une polymérase ; lorsque l'échantillon d'origine est un ARN il convient préalablement d'effectuer une transcription inverse. Il existe actuellement de très nombreux procédés permettant cette 20 amplification, comme par exemple la technique SDA (Strand Displacement Amplification) ou technique d'amplification à déplacement de brin (Walker *et al.*, 1992), la technique TAS (Transcription-based Amplification System) décrite par Kwok *et al.* en 1989, la technique 3SR (Self-Sustained Sequence Replication) décrite 25 par Guatelli *et al.* en 1990, la technique NASBA (Nucleic Acid Sequence Based Amplification) décrite par Kievitis *et al.* en 1991, la technique TMA (Transcription Mediated Amplification), la technique LCR (Ligase Chain Reaction) décrite par Landegren *et al.* en 1988 et perfectionnée par Barany *et al.* en 1991, qui emploie une ligase thermostable, la 30 technique de RCR (Repair Chain Reaction) décrite par Segev en 1992, la technique CPR (Cycling Probe Reaction) décrite par Duck *et al.* en

1990, la technique d'amplification à la Q-béta-réplique décrite par Miele *et al.* en 1983 et perfectionnée notamment par Chu *et al.* en 1986 et Lizardi *et al.* en 1988, puis par Burg *et al.* (1996) ainsi que par Stone *et al.* (1996).

5 Dans le cas où le polynucléotide cible à détecter est un ARN, par exemple un ARNm, on utilisera avantageusement, préalablement à la mise en oeuvre d'une réaction d'amplification à l'aide des amorces selon l'invention ou à la mise en oeuvre d'un procédé de détection à l'aide des sondes de l'invention, une enzyme de type transcriptase inverse afin d'obtenir un ADNc à partir de l'ARN contenu dans l'échantillon biologique. L'ADNc obtenu servira alors de cible pour les amorces ou les sondes mises en oeuvre dans le procédé d'amplification ou de détection selon l'invention.

10 Les sondes nucléotidiques selon l'invention hybrident spécifiquement avec une molécule d'ADN ou d'ARN de polynucléotide selon l'invention, plus particulièrement avec les séquences SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 et SEQ ID N° 13, dans des conditions d'hybridation de forte stringence telles que données sous forme d'exemple précédemment.

15 La technique d'hybridation peut être réalisée de manières diverses (Matthews *et al.*, 1988). La méthode la plus générale consiste à immobiliser l'acide nucléique extrait des cellules de différents tissus ou de cellules en culture sur un support (tel que la nitrocellulose, le nylon, le polystyrène) et à incuber, dans des conditions bien définies, l'acide nucléique cible immobilisé avec la sonde. Après l'hybridation, l'excès de sonde est éliminé et les molécules hybrides formées sont détectées par la méthode appropriée (mesure de la radioactivité, de la fluorescence ou de l'activité enzymatique liée à la sonde).

20 Selon un autre mode de mise en oeuvre des sondes nucléiques selon l'invention, ces dernières peuvent être utilisées comme sonde de capture. Dans ce cas, une sonde, dite « sonde de capture », est

immobilisée sur un support et sert à capturer par hybridation spécifique l'acide nucléique cible obtenu à partir de l'échantillon biologique à tester et l'acide nucléique cible est ensuite détecté grâce à une seconde sonde, dite « sonde de détection », marquée par un élément facilement détectable.

Dans un mode préféré de réalisation, l'invention comprend l'utilisation d'un oligonucléotide sens ou antisens pour contrôler l'expression du produit protéique correspondant. Parmi les fragments d'acides nucléiques intéressants, il faut citer en particulier les oligonucléotides anti-sens, c'est-à-dire dont la structure assure, par hybridation avec la séquence cible, une inhibition de l'expression du produit correspondant. Il faut également citer les oligonucléotides sens qui, par interaction avec des protéines impliquées dans la régulation de l'expression du produit correspondant, induiront soit une inhibition, soit une activation de cette expression. Les oligonucléotides selon l'invention présentent une taille minimale de 9 bases, de préférence de 18 bases, et de manière plus préférée 36 bases.

L'invention concerne un vecteur recombinant de clonage d'un polynucléotide selon l'invention et/ou d'expression d'un polypeptide selon l'invention caractérisé en ce qu'il contient un polynucléotide selon l'invention tel que précédemment décrit. Le vecteur selon l'invention est caractérisé en ce qu'il comporte les éléments permettant l'expression desdites séquences dans une cellule hôte et éventuellement la sécrétion desdites séquences hors de la cellule hôte. Par « vecteur d'expression », on entend aussi bien des vecteurs d'expression à réPLICATION autonome du type plasmide que des systèmes destinés à assurer l'intégration dans les cellules, mais ces vecteurs d'expression pourront être également des vecteurs d'expression de type viral ou bien même, lorsque l'on souhaite réaliser par exemple de la thérapie génique, de l'ADN nu. Parmi les vecteurs viraux, on préfère ceux dérivés de l'adénovirus, du virus associé à

l'adénovirus (AAV), de rétrovirus, des lentivirus, et de préférence les dérivés de HIV, des poxvirus, du virus de l'herpes pour l'expression en système eucaryote. Parmi les vecteurs non viraux, on préfère les polynucléotides nus tels que l'ADN nu ou l'ARN nu selon la technique 5 développée par la société VICAL, les chromosomes artificiels de levure (YAC, yeast artificial chromosome) pour l'expression dans la levure, les chromosomes artificiels de souris (MAC, mouse artificial chromosome) pour l'expression dans les cellules murines et de manière préférée les chromosomes artificiels d'homme (HAC, human artificial chromosome) 10 pour l'expression dans les cellules humaines.

Selon un mode particulier de réalisation, le vecteur selon l'invention comporte des éléments de contrôle de l'expression des polypeptides ; ces éléments de contrôle sont choisis de préférence parmi (i) la séquence promotrice du gène hOBPIIa selon l'invention qui 15 correspond à la séquence SEQ ID N°15 et/ou parmi la séquence promotrice du gène hOBPIIb selon l'invention qui correspondant à la séquence SEQ ID N°16 ; (ii) un polynucléotide dont la séquence est complémentaire à la séquence SEQ ID N° 15 et SEQ ID N° 16 ; (iii) un polynucléotide dont la séquence comporte au moins 80% d'identité 20 avec un polynucléotide défini en (i) ou (ii) ; (iv) un polynucléotide hybride dans des conditions de forte stringence avec la séquence de polynucléotide définie en (i), (ii), (iii). Les outils informatiques à la disposition de l'homme du métier lui permettent aisément d'identifier les boîtes régulatrices promotrices nécessaires et suffisantes au 25 contrôle de l'expression génique, notamment les boîtes TATA, CCAAT, GC, ainsi que les séquences régulatrices stimulatrices (« enhancer ») ou inhibitrices (« silencers ») qui contrôlent en CIS l'expression des gènes selon l'invention ; parmi ces séquences régulatrices, il convient de citer l'IRE, MRE, CRE.

Il est également dans l'étendue de l'invention d'utiliser les éléments ci-dessus définis et choisis parmi la séquence SEQ ID N°15 et/ou SEQ ID N°16 pour contrôler l'expression de polypeptides hétérologues autres que ceux de l'invention et notamment pour diriger 5 l'expression de polypeptides hétérologues dans les types cellulaires dans lesquels les polypeptides selon l'invention s'expriment normalement.

L'invention comprend en outre les cellules hôtes, notamment les cellules eucaryotes et procaryotes, caractérisées en ce qu'elles sont 10 transformées par les vecteurs selon l'invention. De préférence, les cellules hôtes sont transformées dans des conditions permettant l'expression d'un polypeptide recombinant selon l'invention. L'hôte cellulaire peut être choisi parmi les cellules bactériennes mais également parmi les cellules de levure, de même que parmi les cellules 15 végétales et animales ; de préférence, l'hôte cellulaire est une cellule de mammifères (Edwards et Aruffo, 1993), mais également une cellule d'insectes dans laquelle on peut utiliser des procédés mettant en oeuvre des baculovirus par exemple (Luckow, 1993). Ces cellules peuvent être obtenues par l'introduction dans des cellules hôtes d'une 20 séquence nucléotidique insérée dans un vecteur tel que défini ci-dessus, puis la mise en culture desdites cellules dans des conditions permettant la réplication et/ou l'expression de la séquence nucléotidique transfectée.

Selon un mode particulier de réalisation, l'invention porte 25 également sur un animal ou une plante transgénique qui comporte des cellules hôtes selon l'invention.

L'invention concerne également une méthode de préparation d'un polypeptide caractérisé en ce qu'il met en œuvre un vecteur selon l'invention. Plus particulièrement, l'invention porte sur une méthode de 30 préparation d'un polypeptide recombinant caractérisé en ce que l'on cultive des cellules transformées selon l'invention dans des conditions

permettant l'expression dudit polypeptide recombinant et que l'on récupère ledit polypeptide recombinant.

Le polypeptide selon l'invention est susceptible d'être obtenu selon un procédé de l'invention et selon les techniques de production 5 de polypeptides recombinants connues de l'homme du métier. La présente invention concerne donc le polypeptide recombinant susceptible d'être obtenu par la méthode ci-dessus présentée. Dans ce cas, la séquence d'acide nucléique utilisée est placée sous le contrôle de signaux permettant son expression dans un hôte cellulaire. Un 10 système efficace de production d'un polypeptide recombinant nécessite de disposer d'un vecteur, par exemple d'origine plasmidique ou virale, et d'une cellule hôte compatible. Le vecteur doit comporter un promoteur, des signaux d'initiation et de terminaison de la traduction, ainsi que des régions appropriées de régulation de la transcription. Il 15 doit pouvoir être maintenu de façon stable dans la cellule et peut éventuellement posséder des signaux particuliers spécifiant la sécrétion du polypeptide traduit. Ces différents signaux de contrôle sont choisis en fonction de l'hôte cellulaire utilisé. A cet effet, les séquences d'acide nucléique selon l'invention peuvent être insérées 20 dans des vecteurs à réPLICATION autonome au sein de l'hôte choisi, ou des vecteurs intégratifs de l'hôte choisi. De tels vecteurs seront préparés selon les méthodes couramment utilisées par l'homme du métier, et les clones en résultant peuvent être introduits dans un hôte approprié par des méthodes standards telles par exemple la 25 transfection par précipitation au phosphate de calcium, la lipofection, l'électroporation, le choc thermique.

Les polypeptides recombinants obtenus comme indiqué ci-dessus, peuvent aussi bien se présenter sous forme glycosylée que non glycosylée et peuvent présenter ou non la structure tertiaire naturelle.

30 Les polypeptides obtenus par synthèse chimique et pouvant comporter des acides aminés non naturels correspondant auxdits

polypeptides recombinants, sont également compris dans l'invention. Les peptides selon l'invention peuvent également être préparés par les techniques classiques, dans le domaine de la synthèse des peptides. Cette synthèse peut être réalisée en solution homogène ou en phase solide.

Les procédés de purification de polypeptide recombinant utilisés sont connus de l'homme du métier. Le polypeptide recombinant peut être purifié à partir de lysats et extraits cellulaires, du surnageant du milieu de culture, par des méthodes utilisées individuellement ou en combinaison, telles que le fractionnement, les méthodes de chromatographie, les techniques d'immuno-affinité à l'aide d'anticorps mono- ou polyclonaux spécifiques, etc. Une variante préférée consiste à produire un polypeptide recombinant fusionné à une protéine « porteuse » (protéine chimère). L'avantage de ce système est qu'il permet une stabilisation et une diminution de la protéolyse du produit recombinant, une augmentation de la solubilité au cours de la renaturation *in vitro* et/ou une simplification de la purification lorsque le partenaire de fusion possède une affinité pour un ligand spécifique.

La présente invention concerne l'utilisation d'un polypeptide choisi parmi un polypeptide hOBPIIa et hOBPIIb selon l'invention ou l'un de leur fragment, et parmi LCN1, la « rétinol binding protein » (RBP) et l'apolipoprotéine D (ApoD) comme protéine de liaison à un ligand hydrophobe, de préférence une molécule odorante. En effet, outre les polypeptides OBPII mentionnés précédemment, la fonction OBP des lipocalines LCN1, RBP et ApoD n'a jamais été suggérée. En effet, si la présence de LCN1 dans le mucus nasal a été décrite (Redl et al., 1992), c'est en fait une autre fonction qui lui a été attribuée, à savoir une fonction d'anti-protéase, alors que la fonction principale de la protéine LCN1 semble bien être le transport de ligands hydrophobes, d'ailleurs les essais réalisés dans le cadre de la présente invention avec une protéine LCN1 recombinante n'ont pas permis de mettre en évidence

une activité anti-protéase. Il en va de même pour la « rétinol-binding protein » et lapolipoprotéine D qui n'ont jamais été présentées comme ayant une fonction OBP.

La présente invention concerne également l'utilisation d'un polypeptide selon l'invention ou l'un de ses fragments, LCN1, RBP et lapolipoprotéine D comme inhibiteur compétitif, comme agoniste ou antagoniste des récepteurs cellulaires aux lipocalines. L'utilisation d'inhibiteur de liaison des lipocalines à leur récepteur peut être utilisée dans une stratégie anti-cancéreuse ; en effet, un certain nombre de tumeurs sont hormono-dépendantes : ainsi les tumeurs du sein sont sensibles aux stéroïdes ; ces stéroïdes étant transportés par les lipocalines, il est dans l'étendue de l'invention de fournir des inhibiteurs de liaison aux polypeptides selon l'invention et LCN1, OBP ou APO-D pour éviter la fixation des hormones stéroïdes au niveau des récepteurs des cellules tumorales.

L'invention concerne également un anticorps monoclonal ou polyclonal et ses fragments, caractérisés en ce qu'ils lient spécifiquement un polypeptide selon l'invention. Les anticorps chimériques, les anticorps humanisés et les anticorps simple chaîne font également partie de l'invention. Les fragments d'anticorps selon l'invention sont de préférence des fragments Fab ou F(ab')2.

Les polypeptides selon l'invention permettent de préparer des anticorps monoclonaux ou polyclonaux. Les anticorps monoclonaux pourront avantageusement être préparés à partir d'hybridomes selon la technique décrite par Kohler et Milstein en 1975. Les anticorps polyclonaux pourront être préparés, par exemple par immunisation d'un animal, en particulier une souris, avec un polypeptide selon l'invention associé à un adjuvant de la réponse immunitaire, puis purification des anticorps spécifiques contenus dans le sérum des animaux immunisés sur une colonne d'affinité sur laquelle a préalablement été fixé le polypeptide ayant servi d'antigène. Les

anticorps polyclonaux selon l'invention peuvent aussi être préparés par purification sur une colonne d'affinité, sur laquelle a préalablement été immobilisé un polypeptide selon l'invention.

L'invention porte également sur un anticorps monoclonal spécifique d'un polypeptide selon l'invention et capable d'inhiber l'interaction entre ledit polypeptide et le récepteur cellulaire sur lequel se lie spécifiquement ledit polypeptide. Selon un autre mode de réalisation, l'anticorps monoclonal selon l'invention est capable d'inhiber l'interaction entre ledit polypeptide et ses ligands hydrophobes, de préférence les molécules odorantes et de manière préférée les phéromones avec lesquelles ledit polypeptide se lie.

Les anticorps de l'invention pourront également être marqués de la même manière que décrit précédemment pour les sondes nucléiques de l'invention et de manière préférée avec un marquage de type enzymatique, fluorescent ou radioactif. De tels anticorps marqués pourront être utilisés pour la détection de ces polypeptides dans un échantillon biologique. De préférence, l'échantillon biologique est constitué par un fluide par exemple du sérum, du sang ou des biopsies humaines. Ils constituent ainsi un moyen d'analyse de l'expression de polypeptide selon l'invention, par exemple par immunofluorescence, marquage à l'or, immunoconjugués enzymatiques.

Plus généralement, les anticorps de l'invention peuvent être avantageusement mis en oeuvre dans toute situation où l'expression d'un polypeptide selon l'invention doit être observée, et plus particulièrement en immunocytochimie, en immunohistochimie ou dans des expériences de « western blotting », dans les techniques ELISA et RIA. Il est ainsi dans l'étendue de l'invention de fournir une méthode de détection et/ou de dosage d'un polypeptide selon l'invention, dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes de mise en contact de l'échantillon

biologique avec des anticorps selon l'invention puis de mise en évidence du complexe antigène-anticorps formé.

Entre également dans le cadre de l'invention, un nécessaire pour la détection et/ou le dosage d'un polypeptide selon l'invention 5 dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants : (i) un anticorps monoclonal ou polyclonal tel que décrit précédemment ; (ii) le cas échéant, les réactifs pour la constitution du milieu propice à la réaction immunologique ; (iii) les réactifs permettant la détection des complexes antigène-anticorps 10 produits par la réaction immunologique. Ce nécessaire est notamment utile à la réalisation d'expériences de Western Blotting et aux expériences d'immunoprécipitation.

Les hOBPII selon la présente invention peuvent être utilisées dans de nombreuses applications.

15 Les premières applications des OBP résident dans le contrôle des odeurs ; ces applications concernent essentiellement l'hygiène corporelle (parfumerie, cosmétologie, pharmacie) ou collective.

Elles peuvent être utilisées dans un procédé pour contrôler la volatilisation d'une odeur ; un tel procédé se caractérise en ce qu'il 20 comprend une étape de liaison de ladite odeur avec un polypeptide selon l'invention ou avec les protéines LCN1, RBP ou l'apolipoprotéine D.

Il est également possible de fixer l'odeur sur un support solide 25 en utilisant une OBPII selon l'invention attachée audit support, cette fixation pouvant être aussi bien une fixation covalente que non covalente, par exemple par adsorption ou par des fixations de type avidine-biotine, si cela est nécessaire. Dans ces conditions, on obtient un support parfumé retenant les odeurs plus longtemps, libérant progressivement les odeurs et pouvant être utilisé aussi bien dans le 30 domaine cosmétique que dans le domaine des produits d'entretiens en général. Les supports utilisés pourront être des supports de type

plaqué ou de type bille par exemple. Bien entendu, les OBPII selon la présente invention seront plus particulièrement utilisables dans l'industrie de la parfumerie et de la cosmétique où les OBP seront utilisées sous forme d'un mélange liquide destiné à contrôler la volatilisation des odeurs après que la composition ait été répandue sur la peau humaine. Ceci permet de prolonger la « tenue » des parfums ou d'entrer dans la composition des déodorants corporels notamment. Bien entendu, le terme « parfum » englobe de façon générale les mélanges connus en parfumerie, à base d'alcool ou sous forme aqueuse, et contenant notamment des huiles essentielles. Le polypeptide selon l'invention peut également entrer dans la composition de crèmes à base de rétinol en tant qu'agent de transport et de protection dudit rétinol pour des applications cosmétologiques et notamment pour prévenir, effacer et traiter les rides, les ridules de la peau, lutter contre le relâchement cutané et/ou sous-cutané.

Parmi les applications concernant l'hygiène collective, le procédé selon l'invention peut être utilisé dans un dispositif visant à désodoriser les locaux tels par exemple les animaleries, les étables. Un tel dispositif peut également être envisagé pour désodoriser le flux d'air entrant dans un appareil à air conditionné ; un tel dispositif serait très utile dans les zones géographiques polluées.

L'invention concerne également un procédé de criblage de molécule, de préférence les odeurs ou saveurs qui comprend de passer la molécule sur un substrat qui comprend un polypeptide selon l'invention ou, LCN1, RBP et l'apolipoprotéine D lié audit substrat, ledit polypeptide liant ladite odeur ou saveur et de récupérer ladite odeur ou saveur depuis le polypeptide si nécessaire. Les OBP selon l'invention pourraient ainsi permettre l'isolement d'odeurs ou de parfums en les fixant sur un support tel que décrit précédemment mais qui, dans ce cas, pourra être par exemple une colonne de chromatographie, en faisant passer sur celle-ci le produit sur lequel on souhaite capter les

odeurs. Il est bien entendu possible, toujours dans ce même type d'application, d'effectuer une analyse d'un parfum complexe en utilisant les différences de rétention des différents produits par les OBP.

5 Le procédé d'isolement décrit précédemment peut également être utilisé pour séparer des « odeurs » particulières, notamment les phéromones humaines.

Les hOBPII selon la présente invention peuvent être utilisées pour des applications agro-alimentaires ou industrielles.

10 Les OBP selon la présente invention de part leurs caractéristiques peuvent permettre de solubiliser certaines molécules lipophiles en les associant avec lesdits OBP. Ainsi les polypeptides selon l'invention peuvent être utilisés en combinaison avec des acides gras alimentaires à titre d'additif alimentaire. La présente invention 15 concerne un procédé pour solubiliser des molécules lipophiles caractérisé en ce qu'il comprend la liaison de ladite molécule lipophile à un polypeptide selon l'invention, à LCN1, RBP ou à l'ApoD.

Les OBPII, de même que LCN1, RBP et ApoD sont susceptibles d'être impliquées dans le transport des acides gras et dans les 20 mécanismes biologiques permettant la détection de la charge en acides gras de la ration alimentaire, au niveau buccal notamment. L'invention concerne donc également l'utilisation des polypeptides précédents en association avec des acides gras pour diminuer la consommation d'acides gras, notamment dans les hyperlipidémies ou les obésités. Les 25 polypeptides précédents peuvent donc être utilisés pour le traitement des hyperlipidémies et de l'obésité. En effet, ces protéines participent à la détection de la teneur en acides gras dans la prise alimentaire (Gilbertson, 1998), un excès de ces protéines doit conduire à leurrer le système physiologique détectant la charge en gras d'une ration 30 alimentaire. Ainsi, une portion alimentaire pauvre en graisse mais supplémentée en OBP ou ApoD ou RBP ou LCN1 préalablement

chargées en acide gras, sera faussement identifiée comme riche en graisse.

Une autre application est de complémer un lait non maternel avec un ou plusieurs des polypeptides décrits précédemment.

5 Les OBP selon la présente invention peuvent être utilisées en thérapie préventive et curative.

Ainsi, l'absence de détection de ce type de protéines dans un prélèvement biologique en utilisant un anticorps, une amorce, une sonde selon l'invention pourrait être un élément de diagnostic
10 des anosmies.

De même, il est possible de prévoir la détection des anticorps anti-OBPII dans un échantillon biologique d'origine humaine, la présence ou le dosage de ces anticorps pourrait être en rapport avec certains types d'allergies, notamment chez les asthmatiques. Dans ce
15 cas, il pourrait être possible de traiter ce type d'allergie par administration de fragments de polypeptides tels que décrits, ce qui diminuerait les réactions immunitaires aux allergies externes. L'invention concerne donc un procédé de détection d'anticorps dirigé contre les OBPII humaines (hOBPII) dans le sérum humain de patient
20 allergique et/ou asthmatique en utilisant un polypeptide hOBPII selon l'invention.

Selon un autre mode de réalisation, l'invention concerne un procédé de détection d'anticorps dirigé contre les OBPII humaines (hOBPII) dans un fluide biologique de patient atteint d'un cancer, et
25 notamment du cancer de la prostate et/ou du cancer du sein et/ou du cancer de l'utérus et/ou de l'ovaire et/ou du poumon. En effet, les protéines selon l'invention sont sur-exprimées dans les tumeurs et notamment dans le cancer de la prostate (US 5804368, WO 97 10503, CS 84 02898, CS 83 02012, CS 82 08506, CS 82 08215), le cancer du sein (Stoeck et Gould, 1995 ; Simard *et al.*, 1992), le cancer de l'utérus, le cancer de l'ovaire, le cancer du poumon.
30

Les OBP selon la présente invention peuvent également être utilisées dans des compositions pharmaceutiques, ceci notamment afin de vectoriser certaines drogues.

En effet, les lipocalines sont utilisées par les mammifères pour 5 transporter des molécules hydrophobes au sein de fluides biologiques. Il semble même que ce soit les transporteurs naturels des xénobiotiques *in vivo*. Pour exemple, une surcharge expérimentale en xénobiotiques crée chez le rat mâle des tumeurs au niveau du tube contourné proximal (TCP) du néphron (Borghoff *et al.*, 1990). En effet, 10 les protéines MUP, seulement produites chez le rat mâle, sont réabsorbées par les cellules TCP ; après dégradation lysosomiale, les protéines MUP libèrent les xénobiotiques qu'elles transportaient. Ceux-ci s'accumulent dans ces cellules et les conduisent vers une voie tumorale par mutagenèse. Les lipocalines semblent donc être la 15 structure idéale pour piéger en leur calice une molécule utilisée comme médicament, évitant ainsi une distribution générale de celui-ci. Ainsi, G. Beste *et al.*(1999) et la demande WO 99 16873 décrivent une stratégie de mutagenèse et de criblage permettant d'identifier des variants d'une lipocaline ayant une affinité optimale pour un 20 ligand donné. Il est donc possible de créer une « cage à médicament » hautement spécifique.

La présente invention concerne donc un polypeptide selon l'invention en tant qu'agent de ciblage de composé pharmaceutique. Par agent de ciblage, on entend désigner les polypeptides de l'invention 25 capables d'assurer le transport et la libération du ligand auxquels ils sont liés au niveau de certains tissus ou cellules cibles qui possèdent à leur surface des récepteurs auxdits polypeptides. Ainsi, l'invention concerne le transport de médicament au sein de la cage d'un polypeptide selon l'invention et du ciblage de cellules grâce à la 30 capacité de l'OBPII à se fixer à un récepteur spécifique.

Il est également dans l'étendue de l'invention de développer des protéines de fusion ou des protéines chimériques comportant la cage correspondant aux lipocalines selon l'invention associées à une protéine permettant un adressage cellulaire spécifique généralement différent de celui naturel des OBPII. La présente invention a donc pour objet un polypeptide caractérisé en ce que ledit polypeptide est exprimé sous la forme d'une protéine de fusion avec une protéine permettant un adressage cellulaire spécifique. Parmi les protéines permettant un adressage cellulaire spécifique, il convient de citer les interleukines, les cytokines, les lymphokines, les chémokines, les facteurs de croissance, les hormones, les anticorps mono- ou polyclonal. Les interleukines, cytokines et lymphokines sont choisies dans un groupe composé de préférence des interleukines IL-1 à IL-20, des interférons α -IFN, β -IFN et γ -IFN. Les facteurs de croissance sont de préférence les facteurs stimulateurs des colonies (colony stimulating factors) G-CSF, GM-CSF, M-CSF et l'érythropoïétine. Les hormones sont choisies de préférence parmi les hormones stéroïdiennes.

Selon un autre mode de réalisation, le polypeptide selon l'invention en tant qu'agent de ciblage de composé pharmaceutique est associé à une molécule permettant un adressage cellulaire spécifique. Parmi, les molécules permettant un adressage cellulaire spécifique, il convient de citer le groupe des stéroïdes, des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps. De préférence, la molécule est un stéroïde. L'association entre le polypeptide selon l'invention et ladite molécule est réalisée soit par une liaison non covalente en utilisant par exemple le système avidine-biotine ou soit par une liaison covalente en utilisant par exemple des agents chimiques pontants.

Parmi les composés pharmaceutiques que peut transporter et cibler le polypeptide selon l'invention, il convient de citer les médicaments et notamment les agents anticancéreux. Les agents anti-

cancéreux sont sélectionnés parmi le groupe des agents antiprolifératifs, antinéoplastiques ou cytotoxiques et sont utilisés pour arrêter le développement des cancers et pour induire une régression et/ou une élimination de la masse tumorale. Ces agents anticancéreux 5 sont de préférence des radioisotopes, et de manière encore préférée des radioisotopes émetteurs de rayons gamma tels que l'Iode¹³¹, l'Yttrium⁹⁰, l'Or¹⁹⁹, le Palladium¹⁰⁰, le Cuivre⁶⁷, le Bismuth²¹⁷ et l'Antimoine²¹¹. Les radioisotopes émetteurs de rayons beta et alpha peuvent également être utilisés pour la thérapie. Les agents 10 anticancéreux non isotopiques liés au polypeptide selon l'invention sont multiples et variés; on peut citer: (i) les antimétabolites telles les agents anti-folate, le méthotrexate, (ii) les analogues des purines et des pyrimidines (mercaptopurine, fluorouracile, 5-azacytidine), (iii) les antibiotiques, (iv) les lectines (ricine, abrine) et (iv) les toxines 15 bactériennes (toxine diphtérique); les toxines sont choisies de préférence parmi l'exotoxine A de *Pseudomonas*, la toxine diphtérique, la toxine cholérique, la toxine anthrox de *Bacillus*, la toxine Pertussis, la toxine Shiga de *Shigella*, la toxine apparentée à la toxine Shiga, les toxines d'*Escherichia coli*, la colicine A, la d-endotoxine, 20 l'hémagglutinine d'*Haemophilus A*.

L'invention concerne également une composition pharmaceutique comprenant un composé pharmaceutique lié au moins à un polypeptide selon l'invention et un véhicule pharmaceutiquement acceptable. Il est également dans l'étendue de 25 l'invention d'augmenter la capacité de fixation par mutagenèse de cette même protéine. L'invention concerne une composition pharmaceutique telle que définie précédemment pour le traitement du cancer, et notamment le cancer de la prostate, le cancer du sein, le cancer de l'utérus, le cancer de l'ovaire, le cancer du foie et le carcinome des 30 cellules épithéliales pulmonaires. Etant donné que l'OBP dénommée hOBPIIb, l'apolipoprotéine D, la RBP et l'α-1-acide glycoprotéine sont

exprimées dans la prostate, la composition pharmaceutique selon l'invention est destinée au traitement du cancer de la prostate. De même, étant donné qu'il a été démontré que l'ARNm des protéines ApoD, RBP et LCN1 est produit dans la glande mammaire en plus de la 5 production de hOBPIIb déjà décrit dans la demande WO 99 07740, la composition pharmaceutique selon l'invention est destinée au traitement du cancer du sein.

L'invention porte aussi sur un polypeptide selon l'invention en tant que transporteur de composé pharmaceutique. Par transporteur, 10 on entend désigner des polypeptides selon l'invention capable de véhiculer dans l'organisme un composé pharmaceutique sans que ledit composé ne soit libéré à un endroit privilégié dans l'organisme. Un tel polypeptide constitue un moyen de délivrance dudit composé pharmaceutique dans l'organisme.

15 La présente invention concerne également une composition pharmaceutique selon l'invention caractérisé en ce que ledit polypeptide constitue une forme retard de délivrance dudit composé pharmaceutique dans l'organisme.

En effet le composé pharmaceutique d'intérêt lié au polypeptide 20 OBPII selon l'invention diffuse progressivement dans l'organisme à mesure que ledit polypeptide transporteur est catabolisé dans l'organisme. L'utilisation des techniques de l'ADN recombinant permet à l'homme de l'art de modifier la durée de demi-vie du polypeptide OBPII dans l'organisme en introduisant des modifications dans ledit 25 polypeptide. Ainsi, il peut être intéressant de développer un polypeptide homologue selon l'invention pour lequel les sites de coupures protéasiques ont été mutés afin d'augmenter la demi-vie dudit polypeptide dans l'organisme. Il peut être également intéressant de développer un polypeptide multimérique, exprimé par exemple sous 30 la forme d'une protéine de fusion, afin d'éviter une « fuite

glomérulaire » (clearance rénale) et ainsi d'augmenter la demi-vie dudit polypeptide dans l'organisme.

Selon un autre aspect, l'invention concerne un composé caractérisé en ce qu'il est choisi parmi un anticorps, un polypeptide, 5 un ligand, un polynucléotide, un oligonucléotide ou un vecteur selon l'invention à titre de médicament et notamment en tant que principes actifs de médicament ; ces composés seront préférentiellement sous forme soluble, associés à un véhicule pharmaceutiquement acceptable. Par véhicule pharmaceutiquement acceptable, on entend désigner tout 10 type de véhicule employé habituellement dans la préparation de compositions injectables, c'est-à-dire un diluant, un agent de suspension tel une solution saline isotonique ou tamponnée. De préférence, ces composés seront administrés par voie systémique, en particulier par voie intraveineuse, par voie intramusculaire, 15 intradermique ou par voie orale. Leurs modes d'administration, posologies et formes galéniques optimaux peuvent être déterminés selon les critères généralement pris en compte dans l'établissement d'un traitement adapté à un patient comme par exemple l'âge ou le poids corporel du patient, la gravité de son état général, la tolérance au 20 traitement et les effets secondaires constatés, etc.

L'invention concerne également une composition pharmaceutique comprenant un vecteur d'expression selon l'invention et un véhicule pharmaceutiquement acceptable.

Il est également possible d'utiliser la propriété d'expression 25 dans des sites particuliers de ces protéines afin d'assurer une imagerie et/ou un diagnostic de certains types de cancer.

D'autres caractéristiques et avantages de l'invention apparaissent dans la suite de la description avec les exemples et les 30 figures dont les légendes sont représentées ci-après.

FIGURES

FIGURE 1 : Organisation génomique du locus LCN1/hOBPII

La ligne du haut représente la région du chromosome 9q34. La double flèche indique l'intervalle entre la localisation des marqueurs polymorphes D9S1811 et D9S67 ; la position relative des marqueurs D9S67 et D9S1826 est incertaine.

Le niveau du milieu indique l'organisation partielle des cosmides à 3 loci différents : LCN1c, LCN1b-hOBPIIb, LCN1-hOBPIIa. Les 10 cosmides (approximativement 40 kb) sont représentés par des lignes horizontales avec leur nom et les bras des vecteurs T3 ou T7 notés au-dessous de la ligne. Les flèches représentent les différents gènes ou pseudogènes avec leur orientation respective. Dans les cosmides, les lignes en pointillés verticales représentent les sites EcoRI. Le symbole 15 \otimes indique une orientation incertaine du locus.

Le niveau inférieur montre les structures intron/exon des gènes LCN1 et hOBPII : les boîtes noires représentent les exons ; les flèches représentent le site d'initiation de la transcription ; o11 à o15 20 représentent les sondes d'oligonucléotides utilisées pour cibler les cosmides ; *Alu* représente la présence de séquences répétées.

FIGURE 2 : Analyse "Dot Plot" de :

- (A) séquence locus LCN1-hOBPIIa (AC000396+ACXXXXXX) contre locus LCN1b-hOBPIIb (AC002098) ;
- 25 (B) séquence locus LCN1-hOBPIIa (AC000396+ACXXXXXX) contre séquence locus LCN1c ;
- (C) séquence locus LCN1b-hOBPIIb (AC002098) contre séquence locus LCN1c (AC002106).

Des séquences génomiques ont été filtrées pour les séquences 30 répétitives à l'aide de « Repeatmasker » (REF), puis comparées et

« dotplotted » avec le logiciel ggc en utilisant une taille de fenêtre de 25, et un critère de 20 avec une homologie de 80%.

FIGURE 3 : Séquence nucléotidique de gènes hOBPII.

5 Les lignes supérieures représentent la séquence hOBPIIa et les lignes inférieures la séquence hOBPIIb pour laquelle seuls les nucléotides différents sont représentés ; un tiret indique l'absence de séquences correspondantes. Les capitales ombrées sont les séquences exoniques et les lettres minuscules les séquences introniques. Les tailles indiquées sur la gauche sont indiquées en pb. La boîte TATA est en caractères gras et le signal de polyadénylation est souligné. Les boîtes indiquent les sites accepteurs d'épissage pour les exons 5, 5b, 10 5c.

15 **FIGURE 4 : Représentation schématique des deux gènes hOBPII et de leur ARNm correspondants.**

Les lignes horizontales représentent l'organisation exon/intron avec des tailles indiquées en pb. Les boîtes ombrées numérotées de 1 à 20 7 sont les séquences exoniques codantes des principaux transcrits ; les lettres b et c font référence à des exons surnuméraires. Les différents transcrits sont représentés à l'aide de boîtes assemblées : les premières, hOBPIIa α et hOBPIIb α correspondent aux transcrits principaux, les autres correspondent aux formes issues d'un épissage alternatif. ↓ indique un décalage du cadre de lecture résultant de l'insertion ou de la délétion d'un exon et * représente un codon stop. La lettre « a » représente une hélice α et « b » des feuilles β prédis par le logiciel DSC. Les lettres en italique sont des prédictions obtenues avec le logiciel "Predator".

FIGURE 5 : Alignement des séquences des protéines issues des deux gènes hOBPIIa et hOBPIIb humains (hOBPIIa α =

OBP2aaHOMS, hOBPII β α = OPB2baHOMSA, hOBPII β β =
 OPB2bbHOMSA, hOBPII α γ = OBP2agHOMSA, hOBPII α β =
 OBP2abHOMSA), des lipocalines des larmes humaines
 (LCN1_HOMSA), d'OBPII de rat (OBP2_RATNO), de lactoglobuline
 5 bovine BLG (LACB_BOSTA), de MUP de souris (MUP6_MUSMU), de
 RBP humaine (RBP_HOMSA), d'OBP bovine (OBP_BOSTA), de MUP
 de rat (MUP_RATNO), d'OBP porcine (OPB_SUSSC).

Les résidus dans les boîtes gris foncé sont identiques et ceux
 dans les boîtes gris claires sont similaires. Les éléments de structure
 10 secondaire prédis avec le programme DSC sont soulignés et les
 résidus d'amino-acides sont en italiques. Les feuillets β et les hélices α
 sont numérotées pour hOPBIla et b. Le site de clivage prévu du peptide
 signal est indiqué par une flèche (AAA \downarrow LS) en position 15. Des
 séquences non alignées de formes très divergentes de gènes épissés,
 15 hOPBa δ (OPB2adHOMSA) et hOBPby (OBP2bgHOMSA) ont été ajoutées
 au bas de l'alignement après l'analyse.

FIGURE 6 : Analyse RT-PCR

20 (A) Détection des produits RT-PCR LCN1, LCN1b, LCN1c avec
 leurs sondes spécifiques, et (B) Détection des produits RT-PCR
 hOBPIIa et hOBPIIb avec leurs sondes spécifiques. (C) ARN G3PDH
 contrôle. Les tailles sont indiquées en pb.

25

FIGURE 7 : Localisation tissulaire des ARNm hOBP

Des sections du méat moyen (A,B), des cornets (C,D), de la
 prostate (E,F), des canaux déférents (G,H), des glandes mammaires (I,
 J) ont été hybridées à des ribosondes de hOBP marquées à la
 30 digoxigénine-11-UTP. L'hybridation avec une ribosonde hOBP sens n'a
 pas révélé de signal (B,D,F,H, J). Un signal d'hybridation spécifique est
 obtenu avec une ribosonde antisens (A,C,E,G, I). Les flèches indiquent

des structures différentes : AC : cellules acineuses, EC : cellules épithéliales, GC : cellules glandulaires, SD : canal sécrétoire et L : lumen (X200).

5

FIGURE 8 : Arbre de distance phylogénétique des lipocalines des vertébrés.

Les abréviations usuelles des lipocalines ont été employées. Après le symbole "_" sont indiquées les trois premières lettres du genre suivies des deux premières lettres de l'espèce.

10

FIGURE 9 : Schéma de l'évolution de la sous-famille de gène LCN1-OBPII au cours de l'évolution.

15 Les boites indiquent les exons disposés sur une ligne représentant l'ADN génomique. Les « // » entre les lignes indiquent que les loci ne sont pas consécutifs, sans qu'il soit possible de déterminer l'ordre. La large croix sous le symbole « ? » illustre un événement de duplication partielle ou de complète duplication avec une délétion génomique
20 ultérieure pour le locus LCN1c (choix indiqué par le symbole « * »). La croix plus petite indique l'élimination du septième exon qui semble être spécifique de l'homme à cause des nombreuses séquences répétées Alu présentent dans cette région génomique et parce que le rat a deux gènes VEGP. Le symbole « ** » marque l'étape de recrutement d'exons basée sur les données présentes et celles de la littérature ; cette étape a pu apparaître à n'importe quel moment après les duplications des loci et a pu être séquentielle.

25

30 **FIGURE 10 : Analyse par RT-PCR dans la sphère orale et la sphère génitale de l'expression des autres lipocalines humaines connues.**

Définition de nouvelles fonctions OBP et VEGP pour les protéines RBP (retinol-binding-protein) et ApoD (apoliporotéine D).

5 **FIGURE 11 : Etude de l'anticorps polyclonal dirigé contre les protéines hOBPII.**

La protéine de fusion est déposée dans les puits 1 et 3. Un échantillon de mucus nasal humain est déposé dans les puits 2 et 4. Les protéines contenues dans les puits 1 et 2 ont été révélées par coloration au bleu de Coomassie. L'anticorps hybridé sur la membrane obtenue par Western-blotting révèle spécifiquement la protéine GST-hOBPIIb entre 30 et 42 kDa (puit 3) et les hOBPIIa et hOBPIIb (puit 4) de 18 kDa dans le mucus nasal humain.

15

FIGURE 12 : Immunohistochimie sur les tissus de l'appareil olfactif.

Une forte immunoréactivité est repérable sur la coupe de septum (coloration verte, A) sur les deux coupes de cornet (C et D) et sur la coupe de méat moyen (F). Les témoins négatifs réalisés pour le septum (B) et pour le cornet (E) et pour le méat moyen (G) ne montrent aucune réactivité. Les noyaux des cellules sont repérés par une coloration au DAPI (coloration bleue). Les grossissements utilisés sont de X 100 pour les coupes A, B, C, E et G et de X 200 pour les coupes D et F.

25

FIGURE 13 : immunohistochimie sur tissus de la sphère orale

Ces deux tissus, glandes lacrymales (A, B) et glandes de Von Ebner (E, F) montrent une forte immunoréactivité en comparaison des résultats obtenus grâce au sérum pré-immun (C pour les glandes lacrymales et D pour les glandes de Von Ebner). Grossissement X200.

FIGURE 14 : Immunohistochimie sur les glandes mammaires et poumon

- 5 Une forte immunoréactivité est visible sur les coupes de glandes mammaires (A et B) comparé au témoin négatif (C) réalisé grâce au sérum-pré immun. Une forte immunoréactivité est également observée sur les coupes de poumon (E et F) comparée toujours au témoin négatif (D). chaque observation est faite à un grossissement X 100 pour A et C et à un grossissement X 200 pour B, D, E et F.
- 10

EXEMPLES**Exemple 1 : Matériels et Méthodes**5 **1.A. Banque de cosmides du chromosome 9 humain**

Une copie de la banque de cosmides spécifiques du chromosome 9 humain LL09NC01P construite par le Dr. J. Allmeman (Biochemical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA USA) sous l'égide du National Gene Library Project supporté financièrement par le département américain de l'Energie a été utilisée. Le criblage de la banque et l'analyse des clones ont été réalisés tel que décrit précédemment (Lacazette *et al.*, 1997).

15 **1.B. Clonage et analyse de séquence**

Une banque Lambda gt11 d'ADNc de testicules humains (Clontech) (10^7 p.f.u.) a été amplifiée par 30 cycles de polymérisation en chaîne (PCR) (94°C 45 sec, 54 °C 45 sec, 72°C 1min 30 sec) avec l'amorce oliEST58 CCTGCAGGTACATGAGCTTCC et des amplimères 5' ou 3' de criblage d'inserts situés sur les bras du vecteurs Lambda gt11. Une PCR nichée a ensuite été réalisée avec oliEST26 CGCTGTATTGCCAGGCTCC et des oligonucléotides spécifiques du bras du vecteur. Les produits PCR ont été sous-clonés dans le vecteur pGEM-T(r), ce qui a permis d'obtenir l'extrémité 5' des ADNc du gène hOBPII.

25 Les séquences obtenues en utilisant l'oligonucléotide standard pGEM-T (r) et en utilisant un mixe réactionnel prêt à l'emploi de séquençage à base de colorant terminateur (Applied Biosystems) ont été séparées par électrophorèse en utilisant un séquençeur automatique ABI PRISM 377 (Perkin Elmer) puis ont été analysées avec 30 le logiciel Sequence Navigator 1.0.1. (Perkin Elmer). Des clones d'ADNc pleine longueur de hOBPIIa (hOBPIIa α , hOBPIIa β , hOBPIIa δ ,

hOBPII α et hOBPII β (hOBPII α , hOBPII β , hOBPII γ) ont été obtenus à partir de la RT-PCR par purification des bandes d'intérêt selon les instructions du fabricant (gel extraction kit de Qiagen) ou en sous-clonant dans un vecteur pGEM-T(r) les produits de la PCR nichée pour les formes alternatives faiblement exprimées.

1C. Analyse par RT-PCR

Des échantillons de tissus ont été collectés chez des individus caucasiens âgés de 45 à 55 ans en accord avec la réglementation française en vigueur. L'ARN total est extrait selon une méthode en une seule étape utilisant le réactif ARN NOW® selon les instructions du fabricant (Biogentex). 5 µg d'ARN total ont été rétro-transcrits dans un volume final de 20 µl en utilisant 0,5 ng d'oligonucléotide GACTCGAGTCGACATCGATTTTTTTTTTTT avec le système de préamplification Superscript® (Gibco BRL). Trois µl de ces réactions sont ensuite utilisés pour les PCR suivantes. L'expression des ARNm spécifiques a été déterminée par PCR en utilisant : les amorces TL : CCTCTCCCAGCCCCAGCAAG et AP : GACTCGAGTCGACATCG pour les gènes de type LCN1 (LCN1, LCN1b, LCN1c) et pour les gènes de type hOBPII, les amorces DE : CGCCCAGTGACCTGCCGAGGTC, et FI : CTTTATTGGAGTCAGGTGGGTG. Comme contrôle de qualité des ARN, les amorces G3PDH1 : CTCTGCCCCCTCTGCTGATG et G3PDH2 : CCTGCTTCACCACCTCTTG du gène G3PDH ont été utilisées ; le gène G3PDH est considéré comme étant constitutivement exprimé dans tous les types cellulaires. 32 cycles de PCR (94° C 45 sec., 54° C 45 sec., 72° C 2 min. 30 sec.) ont été réalisés et les produits d'amplification ont été séparés sur gel d'agarose à 1%. L'ADN est transféré sur une membrane Hybond N+®.

Pour la détection de l'expression des différents gènes, plusieurs oligonucléotides spécifiques des gènes respectifs ont été synthétisés :

30 - oILCN1 : GACTCAGACTCCGGAGATGA,

- oLCN1b : AACTCAGACACCAGAGATGA,
- oLCN1c : GACTCAGATCCGGAGATGA,
- ol5 : CCAGGAGGGACCACTACA spécifique du gène hOBPIIb,
- ol4 : CCGGGACGGACGACTACG spécifique du gène hOBPIIa,
- 5 - G3PDH3 : CTCATGACCACAGTCCATGC,

Les oligonucléotides sont marqués au $\gamma^{32}\text{P}$ -ATP en utilisant de la T4 kinase (Applied Biosystem) ; les hybridations des oligonucléotides marqués sont réalisées à 42° C. Le lavage final est réalisé dans une solution de 0,1 X SSC, 0,1% SDS à 48° C pendant 20 min. La 10 spécificité des réactions d'hybridation des oligonucléotides est contrôlée en utilisant des échantillons d'ADN cosmique digéré (p233G2 pour LCN1 et hOBPIIa, P19E7 pour LCN1b et hOBPIIb, P181A9 pour LCN1c) et chargés sur le gel avec des produits de RT-PCR.

15

1.D. Génotypage et analyse de liaison

Le génotypage est réalisé par des réactions PCR en utilisant 100 ng d'ADN génomique provenant de 8 familles de référence du CEPH et utilisant les oligonucléotides oli9 TGTCGGGAACGCAGCTT et oli10bis 20 TGCCGCTGTCCCCACGTCGG. Les paramètres du thermocycleur consistent en un cycle initial à 94° C pendant 10 min suivi par 30 cycles à 94° C pendant 30 s, 55° C pendant 30 s et 70° C pendant 45 s puis en une étape d'elongation finale de 10 min à 70° C. Les produits PCR sont ensuite analysés sur un gel d'agarose à 3%. Les informations 25 concernant les marqueurs du chromosome 9 peuvent être obtenues à l'adresse Internet suivante ([HTTP://galton.ucl.ac.uk](http://galton.ucl.ac.uk)) ; les analyses ont été réalisées en utilisant les outils d'études de liaison préalablement décrites dans Lacazette *et al.* (1997). Les haplotypes sont reconstruits manuellement selon les événements de recombinaison préalablement 30 décrits dans la famille 1362 (Attwood *et al.*, 1994).

1.E. Prédictions de structures secondaires

Un alignement multiple des protéines lipocalines pour lesquelles les structures cristallographiques ont déjà été décrites (Monaco *et al.*, 1992 ; Spinelli *et al.*, 1998) avec les protéines hOBPIIa et hOBPIIb a été obtenu en utilisant le logiciel Clustalx (<ftp://ftp.inforbiogen.fr>). Les structures secondaires putatives ont été déterminées avec le programme DSC (Discrimination of protein Secondary structure Class) développé par R.D. King et M.J.E. Sternberg (<http://bioweb.pasteur.fr/seqanal/interfaces/dsc-simple.html>). Les structures secondaires des protéines correspondantes aux formes alternativement épissées sont supposées identiques aux formes traditionnelles avant le décalage du cadre de lecture; après ce décalage la prédition de structure est effectuée avec une seule séquence et est réalisée en utilisant le logiciel Predator (<http://pbil.ibcp.fr/cgi-bin>).

15

1.F. L'analyse phylogénétique

Les protéines lipocalines dont les séquences sont connues dans leur totalité ont été alignées trois fois consécutivement en utilisant le logiciel Clustalx (<ftp://ftp.inforbiogen.fr>). Des distances dans l'arbre phylogénétique ont été calculées avec Clustalx et dessinées avec Njplot.

1.G. Hybridation *in situ*

Des coupes séries réalisées au cryostat (8 µm d'épaisseur) sont collectées sur des lames SuperFrost® Plus (Menzel Glazer) et stockées à - 80° C. Des sondes ARN antisens et sens sont obtenues selon les techniques standard en utilisant la T7 ou la SP6 polymérase à partir d'une matrice obtenue par digestion avec des enzymes de restriction NcoI ou PstI du clone cDNA de phOBPIIaP2 et en utilisant du DIG-11-UTP (Boehringer Mannheim) (la longueur de la sonde est approximativement de 150 nucléotides). Les matrices digérées par PstI

et transcrives par l'ARN polymérase de T7 correspondent à la sonde antisens et les matrices digérées par NcoI transcrives par l'ARN polymérase SP6 correspondent à la sonde sens. Des sections de tissus sont fixées dans du paraformaldéhyde 4% pendant 15 min puis rincées 5 pendant 5 min dans du PBS 2X. Après acétylation (2 x 5 min dans un tampon de triéthanolamine (TEA) pH 8, contenant 0,25% v/v d'anhydride acétique), les sections tissulaires sont préhybridées à 60° C pendant 15 min dans de la formamide 50%/1X SSC. Les sondes marquées sont appliquées sur chacune des sections dans 50 µl de 10 tampon d'hybridation (50% formamide, 1X Denhardt's, 500 µg/ml total d'ARNt, 10% de Dextran sulfate, 10 mM de dithiotréitol). Les sections sont recouvertes puis incubées dans des chambres humides à 50° C pendant la nuit. Après hybridation, les lames sont immergées à 55° C dans un tampon de lavage (50% formamide, 1X SSC) pendant 2 15 heures. Chaque lame est ensuite rincée 2 fois 5 min dans du 2X SSC à température ambiante, puis traitée pendant 30 min avec 10 mg/ml de RNase à 37° C, et enfin immergée 2 heures à 55° C dans une solution de lavage (50% formamide, 2X SSC). Les lames sont ensuite placées pendant 15 min à 55° C dans du 0,1 X SSC. La détection 20 immunologique est réalisée en utilisant un anticorps anti-DIG conjugué à la phosphatase alkaline (fragments Fab) selon le protocole de Boehringer Mannheim. Les sections sont examinées à des grossissements différents utilisant un microscope Axiophot (Zeiss).

25 **1.H. Obtention d'un anticorps polyclonal dirigé contre les protéines hOBP_{II}**

Deux cent cinquante µl de protéines de fusion ont été fournis à la société Agro-Bio pour la production d'anticorps. Des lapins (New Zealand White SPF) sont injectés aux jours 0, 14, 28 et 42. Les serums 30 sont prélevés aux jours 0, 35, 49 et 63. Afin de tester l'anticorps, la protéine de fusion est déposée dans les pistes 1 et 3 (Figure 11), alors

que du mucus nasal humain est déposé dans les pistes 2 et 4. Les pistes 1 et 2 ainsi que la piste correspondant à l'échelle de poids moléculaire, correspondent à une révélation par le bleu de Coomassie. Les pistes 3 et 4 correspondent à une détection des protéines 5 reconnues par l'anticorps à l'aide de la technique de Western blot. Les pistes 1 et 3 montrent la présence d'une série de protéines recombinantes tronquées entre 30 et 42 kDa issues d'une synthèse peu efficace due à la présence de nombreux codons rares chez la bactérie au sein de la séquence de hOBPIIb. Toutefois cette production 10 de protéine recombinante fût suffisante pour produire un anticorps polyclonal de bonne qualité comme indiqué par la révélation dans la piste 4 correspondant au mucus nasal, d'une bande de 18 kDa spécifique des protéines hOBPII.

15 **1.I. immunohistochimie**

Des coupes sériées de 8 µm de différents tissus sont fixées au para-formaldéhyde (5min), et rincées trois fois au PBS x1 (15 min) puis incubées 30 min dans une solution 3% de BSA en PBSx1, avant d'être incubées durant la nuit en présence de l'anticorps anti-protéine 20 de fusion en solution PBSx1. Après trois rinçage en PBSx1 (15 min), un anticorps anti-IgG de lapin couplé au FITC (coloration verte en fluorescence) en solution de PBSx1 est placé 3 heures au contact des lames. Après un rinçage en PBSx1, une solution de DAPI (100 ng / µl en PBSx1) est appliquée pendant 10 min (contre coloration des noyaux 25 cellulaires en bleu). Après trois rinçages en PBSx1 (15 min), les lames sont montées dans de l'eau glycérinée (50/50). L'analyse est réalisée en présence de filtres DAPI et FITC à l'aide d'une caméra CDD en utilisant un temps d'intégration entre 4 et 32 ms.

Exemple 2 : Identification d'un gène homologue à LCN1 localisé sur le chromosome 9 humain

L'identification de l'ADN complémentaire de LCN1 codant pour la lipocaline des larmes humaines (Lassagne et Gachon, 1993) a préalablement été rapportée, ainsi que la localisation du gène sur le chromosome 9q34 (Lassagne *et al*, 1993 ; Lacazette *et al*, 1997). Les deux gènes codant pour les protéines des glandes de von Ebner 1 et 2 (Kock *et al*, 1994) correspondent aux protéines homologues de LCN1 chez le rat ; ceci pose la question de savoir s'il existe d'autres gènes codant pour une protéine LCN1 dans le génome humain. Les expériences d'hybridation *in situ* (Lassagne *et al*, 1993) ainsi que les analyses avec des hybrides somatiques (Lassagne *et al*, 1995) indiquent que, s'ils existent, ces gènes humains additionnels doivent se trouver localisés dans la région chromosomique 9q34.

Une banque de cosmides spécifiques du chromosome 9 humain (LL09NP01) a été criblée avec la sonde d'ADNc LCN1 humaine ; 26 cosmides ont été identifiés ; ceux-ci ont été digérés par EcoRI ou Pvull, puis hybridés successivement avec l'ADN de LCN1 et différents oligonucléotides (figure 1).

Les cosmides sont divisés en 3 groupes. Le premier groupe (clones P32H3, P41B5, P63B6, P92H20, P109C6, P145H6, P195B4, P233G2, P233F2, P265D4, et P276H8) correspond aux cosmides contenant une séquence du gène LCN1 préalablement isolé (numéro d'accession : L14927) formé par 7 exons (Holzfeind et Redl, 1994). Le second groupe (clone P19E4, P19E7, P42H9, P98H5 et P142H8) correspond à la séquence LCN1b homologue à LCN1 (clone P19E4) (numéro d'accession : Y10826) depuis le promoteur jusqu'au 6^{ème} exon puis qui diverge ensuite. Un troisième groupe de cosmides (clone P110C1, P174E4, P174E5, P181A9, P181B6, P211A7, P238C6 et P291E1) contient une région LCN1c, établi à partir du séquençage partiel du clone P181A9 (numéro d'accession Y10827), qui est fortement

homologue à LCN1 seulement à partir du promoteur jusqu'à l'exon 2. Ainsi, LCN1 est le seul gène qui possède le 7^{ème} et dernier exon. De plus, les boîtes TATA sont dégénérées dans les promoteurs des gènes LCN1b et LCN1c.

5

Exemple 3 : Identification de duplications génomiques contenant les gènes de lipocalines et cartographie de la région chromosomique 9q34

10 Au cours de la période d'identification de la famille de gène LCN1, un vaste projet de cartographie physique a conduit à l'identification de contigs de cosmides couvrant partiellement la région chromosomique 9q34 (Nahmias *et al.*, 1994, Van Slegtenhorst *et al.*, 1995, Hornigold *et al.*, 1997) ; le séquençage des clones correspondants a été réalisé par
15 un groupe du Massachusset Institute of Technology (Boston, USA).

Des analyses de comparaison de séquences dans les banques de données de LCN1, LCN1b et LCN1c ont révélé des homologies fortes entre le gène LCN1 préalablement décrit (numéro d'accession L14927) et les trois séquences cosmides : P161A1 ayant le numéro d'accession
20 AC002098, P203H12 avec le numéro d'accession AC000396 et P161G2 avec le numéro d'accession AC002106.

Une analyse très fine notamment de la région 3' des gènes LCN1 a révélé que la séquence LCN1c correspondait au clone ayant le numéro d'accession AC002106, LCN1b au numéro AC002098 et LCN1 au numéro AC000396 ; il existe cependant une insertion de 60 paires de base à la position 12360 du clone ayant le numéro d'accession AC000396 par rapport à celui ayant le numéro L14927. Des analyses complémentaires ont révélé que ces séquences sont homologues sur une région beaucoup plus large que les gènes LCN1. Comme illustré
25 par les analyses en dot plot (figure 2), ces trois cosmides correspondent à des régions génomiques dupliquées ; les cosmides P161A1
30

(AC002098) et P203H12 (AC000396) sont homologues sur la totalité de leurs séquences ; le cosmid P161G2 (AC002106) n'est homologue à la séquence uniquement pour des séquences situés en amont de l'intron 3 du gène LCN1.

- 5 Les positions relatives de ces régions dupliquées du chromosome 9 humain ont été analysées. Les résultats ont été comparés avec la carte physique établie dans le cadre du projet de cartographie physique par des analyses de restriction des cosmides (Hornigold *et al.*, 1997) en utilisant la même banque LLN09NC01P. Approximativement 10 les mêmes cosmides ont été trouvés dans les deux recherches pour le locus LCN1c ; la comparaison de séquence du cosmid P181A9 révèle que son extrémité T3 contient le gène Surf 5 (figure 1). De plus, AC002106 est localisé dans un contig de séquence entre ABO et le locus Surfeit confirmant la localisation de LCN1c. P161A1 et P203H12 15 ont des localisations moins précises en utilisant les données de Hornigold et de ses collaborateurs à cause de la limite de la stratégie de cartographie de restriction pour les régions dupliquées. P203H12 (AC000396) semble correspondre au gène LCN1 qui a été préalablement localisé près du marqueur D9S1826 (Lacazette *et al.*, 20 1997) à l'exception d'une insertion de 60 paires de bases.

Pour tester si P203H12 peut correspondre à une 4^{ème} région dupliquée, la région a été testée par PCR sur de l'ADN génomique de 150 individus non apparentés en utilisant des amorces spécifiques pour AC000396 et L14927. Une bande unique correspondant à la 25 longueur de la séquence AC000396 a été détectée (donnée non montrée) prouvant ainsi que P203H12 contient le gène LCN1 préalablement décrit. Le locus LCN1b n'a pas été positionné précisément dans la région comme il est démontré par le fait que les extrémités AC002098 ainsi que les extrémités des cosmides contenant 30 LCN1b ne détectent aucune séquence homologue dans les banques de données.

L'analyse des séquences a permis d'identifier, un mini-satellite putatif à la position 3177-3724 du cosmide correspondant à AC002098 (Figure 1) ; un polymorphisme rare a ainsi été mis en évidence dans une population de 20 individus non apparentés (PIC = 0,05). Ce nouveau marqueur polymorphe n'est informatif que dans la seule famille 1362 sur les 8 familles de référence du CEPH testées.

L'analyse de liaison a révélé des lod-scores à deux points supérieurs à 3 à téta = 0 pour les marqueurs D9S275 et D9S1818. La reconstruction de l'haplotype a confirmé la localisation du gène LCN1b entre les marqueurs D9S1811 et D9S67 dans le chromosome 9q34 (Figure 1).

Exemple 4 : Identification de deux nouveaux gènes de lipocaline

Les analyses de séquences ont révélé des données nouvelles.

La comparaison de AC002098 à la banque de données a révélé des similitudes entre cette séquence et les gènes de lipocalines. En plus de la région 21000 à 27000 qui contient le gène LCN1b, la région autour de la position de 2150 contient des séquences présentant des similitudes avec les séquences codant pour les protéines de liaison aux odeurs de rat de type II, des lipocalines des larmes ainsi que pour l'EST AA460385.

Parallèlement, la comparaison de la séquence AC000396 à la banque de données a révélé en plus de la région 11100 à 17100 contenant le gène LCN1, des similitudes pour le même groupe de séquences dans la région 36600 à 37800.

L'EST AA460385 exprimée dans les testicules humains correspond à 4 exons du nouveau gène de lipocaline présent dans le clone P161A1 (AC002098). Un exon putatif de 50 paires de bases similaire à la séquence EST est également présent à l'extrémité du clone P203H12 (AC000396).

Les inventeurs ont donc conclu à l'existence d'un nouveau gène de lipocaline à une position distale de LCN1b. Les inventeurs ont émis l'hypothèse, que suite à une duplication génomique, un second gène orthologue au gène codant pour l'EST est présent dans la région 3' de 5 LCN1.

Le séquençage du cosmid P233G2 contenant cette région (Figure 1) avec des oligonucléotides définis à partir d'EST a en effet révélé la présence d'un autre gène de lipocaline à un locus distant de 10 20 kb distal du gène de LCN1. Pour identifier les premiers exons des deux nouveaux gènes, des PCR nichées sur des clones cDNA provenant d'une banque de testicules ont été réalisées entre les oligonucléotides localisés dans la région 5' de l'EST et les bras du vecteur. Les produits PCR sont clonés et leur séquence a révélé trois exons additionnels d'après la séquence génomique (Figure 1). Une boîte TATA est présente 15 en amont du premier exon dans les deux cas (Figure 3). Les deux ARNm, hOBPIIa correspondant au gène localisé en aval de LCN1 et hOBPIIb localisé en aval de LCN1b (Figures 3 et 4), sont identiques à 97,5% et 63 % à ceux correspondant au gène LCN1. Les organisations 20 intron/exon de ces deux gènes sont consistantes avec la famille des lipocalines.

De plus, les séquences protéiques déduites (170 amino acides) confirment l'appartenance à la famille des lipocalines. Les deux protéines matures putatives hOBPIIa (MW = 17,8 kDa) et hOBPIIb (MW = 18,0 kDa) sont identiques à 89%. Ces deux nouvelles protéines possèdent un peptide signal putatif de 15 amino acides (Figure 5), les résidus conservés GXWY à la position 27 à 30 et deux cystéines susceptibles d'être impliquées dans un pont disulfure (position 74 à 25 166).

Cependant, les comparaisons de séquences par paire avec les autres lipocalines indiquent une faible conservation des résidus acides aminés (15 à 25%) à l'exception de TL/VEG humain codé par le gène 30

LCN1 (valeur moyenne 43 % sur 155 amonoacides) et l'OBPII de rat (valeur moyenne 45,5 %). Par ailleurs, les points isoélectriques calculés de hOBPIIa et de hOBPIIb sont respectivement de 7,85 et de 8,72 alors que ceux des lipocalines sont acides (aux environs de 4,5) à l'exception de l'OBPII de rat (PI = 9,01). La prédition des structures secondaires des protéines hOBPIIa et hOBPIIb avec le logiciel DSC en effectuant des alignements multiples avec les séquences de lipocalines de structures connues indique la présence de 8 brins de feuillets β antiparallèles pouvant permettre la formation d'un calice suivi par une hélice α et un feuillet β final en accord avec les données connues des structures des lipocalines (Figure 5).

Exemple 5 : Etude de l'expression des gènes de lipocalines identifiés

Pour préciser si LCN1b, LCN1c, hOBPIIa et hOBPIIb sont des gènes exprimés, des analyses de RT-PCR ont été réalisées sur 18 tissus humains différents connus pour produire des lipocalines (Figure 6).

Deux couples d'amorces ont été synthétisés ; l'un reconnaît les ARNm de type LCN1 et l'autre les ARNm de type hOBPII. L'hybridation des oligonucléotides spécifiques des gènes sur des membranes sur lesquelles sont immobilisés les produits de RT-PCR après transfert par la technique de Southern, puis le sous-clonage des produits hybridés ont révélés que LCN1b et LCN1c ne sont exprimés dans aucun des produits testés, alors que les ARNm de LCN1 sont détectés dans la glande lacrymale, les glandes sudoripares, les glandes de von Ebner, le septum nasal, l'épithélium du cornet nasal, tout comme le placenta et les glandes mammaires (Figure 6A). Ces données et les informations de séquences qui indiquent que la boîte TATA et le dernier exon ont été perdus permettent d'affirmer que LCN1b et LCN1c sont des

pseudogènes qui ne participent pas à la formation des protéines LCN1 humaines.

A l'inverse, les gènes hOBPIIa et hOBPIIb sont exprimés, ce qui confirme leur détection précédente dans les banques d'ADNc. De manière surprenante, bien que les deux protéines hOBPIIa et hOBPIIb soient très similaires sur la totalité de leur séquence y compris dans la région promotrice de 1,5 kb, leur profil d'expression sont différents (Figure 6B). La protéine hOBPIIa est fortement exprimée dans le septum nasal, le méat moyen, le cornet nasal, les testicules et le placenta et plus faiblement dans les glandes mammaires, les glandes lacrymales, les glandes sudoripares, les glandes de von Ebner et le poumon. A l'inverse, la protéine hOBPIIb est exprimée de manière prédominante dans la prostate, les testicules et les glandes mammaires et plus faiblement dans les glandes sous-maxillaires, le septum nasal et le méat moyen.

De plus, les analyses de RT-PCR ont révélé l'existence d'un épissage alternatif du produit de transcription du gène hOBPIIa et du gène hOBPIIb qui génère respectivement quatre et trois ARNm (Figures 3, 4, 5 et 6).

La transcription du gène hOBPIIa génère au moins quatre ARNm qui codent pour quatre protéines différentes; le premier ARNm code pour la protéine hOBPIIa α qui correspond à la protéine hOBPIIa décrite précédemment. Dans le gène hOBPIIa, trois sites accepteur d'épissage différents ont été identifiés pour l'exon 5 (figure 3 et 4) formant ainsi deux autres variants d'épissage. Un premier variant d'épissage présente un site accepteur pour l'exon 5 localisé 49 pb avant le précédent (exon 5b) ; ceci génère un ARNm de 725 nucléotides qui code pour la protéine hOBPIIa β de 146 aminoacides. Cette protéine est identique jusqu'au 8^{ème} feuillet β putatif puis différente ensuite avec seulement 16 aminoacides additionnels. Un second variant d'épissage présente un site accepteur pour l'exon 5 localisé 65 pb avant le

5 précédent (exon 5c) ; ceci génère un ARNm de 741 nucléotides qui code pour une protéine hOBPII α y de 228 aminoacides. Cette protéine possède les huit premiers feuillets β putatifs identiques à ceux de hOBPII α puis est différente dans la région C-terminale (figure 5) à cause d'un décalage de cadre de lecture généré par cet événement d'épissage alternatif ; la structure de cette région C-terminale prédictive par le logiciel Predator est une longue région coudée contenant une 9^{ème} feuillet β .

10 Dans le cas du gène hOBPIIb, en plus de l'ARNm hOBPII α préalablement décrit, un exon surnuméraire de 106 pb (exon 3b) entre les précédents exons 3 et 4 a été identifié (figure 3). Cet ARNm plus long (782 nucléotides) code pour une protéine hOBPII β b de 165 amino-acides. Du point de vue de la structure protéique, hOBPII β b est identique à hOBPII α jusqu'au 5^{ème} feuillet β putatif puis diffère ensuite à cause d'un décalage du cadre de lecture. Les prédictions des logiciels informatiques indiquent que le motif ALWEALAIIDTRLK est une hélice α qui est juste derrière le cinquième feuillet β . Deux feuillets β additionnels peuvent être présents dans la longue partie C-terminale.

20 Aucun des variants d'épissage alternatifs pour un gène n'est détecté symétriquement pour l'autre gène bien que les sites d'épissage accepteur et donneur putatifs puissent être présents (figure 3). En plus de ces motifs protéiques qui conservent la structure traditionnelle d'une lipocaline, nous avons démontré, pour les deux gènes hOBPII α et hOBPIIb, l'existence d'une faible quantité d'ARNm ayant subi un épissage alternatif et codant pour des protéines dont la séquence n'est pas directement liée aux lipocalines. Des ARNm codant pour hOBPII α d et hOBPII β y auxquels ils manquent la séquence codant pour l'exon 2 et qui possèdent respectivement l'exon 5b et l'exon 5 codent pour des protéines putatives sécrétées de 147 et de 85 amino-acides respectivement (figures 4 et 5) ; ces protéines divergent des protéines précédentes à partir du 24^{ème} amino-acide.

Pour identifier les cellules exprimant les gènes hOBPII, les inventeurs ont hybride des sondes ribonucléiques sens et anti-sens marquées à la digoxygénine sur des sections tissulaires (figure 7). Les ARNm hOBPII sont détectés dans les cellules acineuses du méat moyen et des cornets nasaux ainsi que des cellules épithéliales des cornets ; ceci soutient l'idée que les protéines hOBPII sont impliquées dans la fonction olfactive. En plus de la production dans la sphère orale, des ARNm codant pour hOBPII ont été détectés dans la sphère génitale, notamment dans les cellules glandulaires de la prostate, dans les cellules secrétrices épithéliales du canal déférent. Aucun signal n'a été détecté dans les gonades mâles, ce qui suggère que l'expression des gènes hOBPII mise en évidence dans les expériences de RT-PCR correspondait à la présence de canaux additionnels dans la préparation tissulaire (rete testis et canaux efférents). En combinant ces résultats de la détection de l'ensemble des ARNm codant pour hOBPII avec ceux de l'approche RT-PCR, il est apparu que cinq protéines hOBPII (hOBPII α , hOBPII β , hOBPII γ , hOBPII δ , hOBPII ϵ) sont sécrétées par les cellules épithéliales des canaux des gonades mâles, ainsi que des cellules acineuses du méat moyen et des cornets nasaux ; dans ces cellules les ARNm codant pour hOBPII α sont hautement prédominant. Seules les deux protéines hOBPII δ (hOBPII $\delta\alpha$ et hOBPII $\delta\beta$) sont sécrétées par les cellules épithélio-glandulaires de la prostate et des glandes mammaires.

25

Exemple 6 : Analyse phylogénétique et classification des lipocalines

L'analyse en dotplot (figure 2) ainsi que les duplications génomiques que nous avons révélées indiquent une origine commune des gènes LCN1 et hOBPII. Pour clarifier les relations entre les membres de la famille des lipocalines et pour vérifier que nous avons

identifié le gène humain orthologue de OBPII de rat (Dear *et al.*, 1991), nous avons construit un arbre de distance phylogénétique avec les lipocalines des vertébrés (figure 8).

Les inventeurs ont mis en évidence neuf groupes principaux de
5 lipocalines issus d'un précurseur commun :

- la famille de l'apolipoprotéine et de la protéine de liaison au rétinol (RBP) (groupe 1) ;
- le groupe de la prostaglandine D-synthase et du précurseur de lipocaline associé à la gélatinase des neutrophiles (groupe 10 2) ;
- la sous-famille de l'alpha-1-microglobuline/bikunin (protéine HC) (groupe 3) ;
- la sous-famille orosomucoïde (A1AG, A1AH, A1AI) (groupe 4) ;
- la sous-famille de la sphère orale 1 (OBPII-type-LCN1/VEGP, VNSP I et II, LALP, CanF1) (groupe 15 5) ;
- la sous-famille de la lactoglobuline (groupe 6) ;
- la sous-famille des protéines sécrétées par l'épididyme de lézard (groupe 7) ;
- la sous-famille la sphère orale 2 (protéines urinaires majeures (MUP) de souris) (groupe 20 8) ;
- la sous-famille de la sphère orale 3 (OBP1, OBPII de souris, Aphrodisine, Probasine, BD20) (groupe 9).

La présente invention concerne plus particulièrement les groupes 5, 8 et 9.

25 Le groupe 5 contient les protéines hOBPII qui sont étroitement liées d'un point de vue évolutif à l'OBPII de rat. Pris dans leur ensemble, les présents résultats indiquent que les gènes hOBPII humains sont orthologues du gène OBPII de rat. Ce groupe contient également les protéines LCN1-VEGP de différentes espèces. En prenant 30 en considération l'organisation génomique des gènes hOBPII-LCN1, les données de l'arbre illustrent l'événement de duplication (flèche) qui

donne naissance aux gènes ancestraux hOBPII et LCN1 à partir de leur précurseur lipocaline commun (Figure 9). Plus récemment, les duplications originelles de la région de 50 kb contenant hOBPII-LCN1 (flèche) ont généré chez l'homme les deux gènes hOBPII et le gène 5 LCN1 et son pseudogène LCN1b. La duplication additionnelle ayant donné lieu au pseudogène LCN1c est partielle dans le génome humain et ne produit pas de protéine fonctionnelle et est donc absente du présent arbre. De plus, les deux protéines VEG du rat sont plus étroitement liées l'une à l'autre dans l'arbre phylogénétique qu'avec la 10 protéine humaine LCN1. La situation est identique pour les deux protéines humaines OBPII par rapport à la protéine OBPII de rat. Ces résultats sont en faveur d'un processus de conversion génique pour au moins certains gènes des lipocalines. Ceci peut également être corrélé au fait que ces gènes se situent sur le même bras chromosomique.

15 Le groupe 9 correspond à la famille de la sphère orale 3 et contient des OBP qui ont déjà été décrites. La protéine Aphrodisine a déjà été décrite comme un transporteur de phéromone (Henzel *et al.*, 1988) et apparaît être orthologue à la protéine OBPI de rat et de souris, avec deux gènes OBPI paralogues. Finalement, les lipocalines 20 olfactives produites par les glandes de Bowman de *rana pipens* (OLFA RANPI) qui sont considérées comme des transporteurs d'odeurs potentiels dans le mucus de la grenouille, ne sont liées à aucun groupe d'OBP putatif (groupes 5, 8 et 9) suggérant ainsi l'existence d'autres catégories d'OBP.

25 Certaines lipocalines ont été décrites comme allergènes. L'allergène majeur du chien, la protéine CanF1 majoritairement exprimée dans les glandes de von Ebner (Konieczny *et al.*, 1997), apparaît dans l'arbre être la protéine VEG du chien. Ce résultat montrant que LCN1 peut être allergénique, pousse les inventeurs à 30 proposer l'implication des protéines OBPII dans les processus allergiques. De la même manière, l'allergène majeur du cheval EquC1

exprimé dans le foie, dans les glandes salivaires (Grégoire *et al.*, 1996) apparaît être orthologue à un membre de la protéine MUP (groupe 3). L'allergène majeure de la vache (BD20 BOSTA) qui est présent dans la famille de la sphère orale 2 (groupe 4) est davantage lié à la probasine.
5 CanF2 principalement exprimée dans la glande parotide (Konieczny *et al.*, 1997) ne semble pas être orthologue à une lipocaline préalablement décrite.

La présente invention a permis de révélé l'existence d'une duplication génomique au locus q34 du chromosome 9 humain qui recèle une famille de gène du type LCN1 ; cette famille comporte autre le gène LCN1, précédemment décrit, deux pseudogènes ainsi que deux gènes hOBPII qui sont paralogues à LCN1. Les inventeurs ont révélé que la famille hOBPII-LCN1 résulte d'événements consécutifs de 10 duplications génomiques. Les séquences ainsi que l'organisation génomique ont révélé que les gènes LCN1 et hOBPII dérivent d'un ancêtre commun et ont été générés au moyen de duplication en tandem. 15 Les comparaisons de séquences ont démontré que les protéines hOBPIIa α et hOBPIIb α sont les formes de protéines des hOBPII les plus proches de LCN1; ceci est corroboré par le fait que la taille de l'exon 5 de LCN1 est de 121 pb, ce qui correspond à la taille de l'exon 5a de la protéine hOBPII et par le fait qu'aucun exon 3b n'a été trouvé dans la protéine LCN1. Des résultats similaires sont obtenus lorsque cette 20 comparaison est effectuée avec les sept autres exons des lipocalines. Ces données sont en faveur d'une acquisition de la diversité des protéines hOBPII à travers l'intégration de l'ADN génomique additionnel environnant au niveau du site accepteur d'épissage amont pour l'exon 5 de hOBPIIa ou à travers le recrutement d'un exon 25 surnuméraire (exon 3b) pour la protéine hOBPIIb. L'hypothèse d'un recrutement d'exons plutôt qu'une réduction de la taille des exons ou 30 qu'une perte d'exons est plus probable.

Les inventeurs ont montré que les gènes hOBPII et LCN1 codent pour des protéines impliquées dans différentes fonctions, comme le montre l'expression de ces gènes à la fois dans la sphère orale et dans la sphère génitale. De plus, les inventeurs ont montré par l'analyse phylogénétique (exemple 6) que plusieurs protéines différentes pouvaient participer à la même fonction d'odorant-binding protéine ; ainsi trois sous-familles de lipocalines (les groupes 5, 8 et 9) correspondant à plusieurs protéines sont retrouvées exprimées dans la sphère orale et notamment dans les glandes nasales et buccales.

Cette analyse montrant que plusieurs lipocalines participent à une même fonction physiologique et qu'une même protéine pouvait participer à différentes fonctions, a conduit les inventeurs à analyser l'expression des autres lipocalines humaines dans l'ensemble des tissus étudiés (Figure 10). Ceci a conduit à montrer que le gène codant pour l'apolipoprotéine D est exprimé dans les glandes du cornet nasal et du méat moyen, en faisant ainsi une potentielle odorant-binding protéine. De même la rétinol-binding protéine (RBP) est exprimée par l'épithélium nasal et pourrait aussi participer à cette fonction.

Les inventeurs proposent donc d'inclure les protéines ApoD et RBP dans la famille des OBP humaines et les incluent dans l'ensemble des revendications portant sur ces OBP humaines.

Les inventeurs ont démontré que la transcription des deux gènes hOBPII génère de nombreux transcrits alternatifs qui codent pour des protéines distinctes dont la structure est compatible avec celle d'un transporteur de ligand hydrophobe.

Les inventeurs ont mis en évidence l'expression des deux gènes hOBPII dans la sphère orale (glandes nasales, glandes de von Ebner, glandes sous-maxillaires, glandes lacrymales, poumon). Les inventeurs ont également démontré que les protéines hOBPII selon l'invention sont produites par les cellules de la sphère génitale ; le gène hOBPII est majoritairement exprimé dans la prostate, le canal déférent et les

glandes mammaires alors que l'expression du gène hOBPIIa est restreinte au canal déférent. Les analyses *in situ* ont révélé que l'ARNm est produit par les cellules glandulaires de la prostate et des cellules sécrétrices épithéliales du canal déférent supportant l'idée d'une sécrétion des protéines correspondantes dans le fluide séminal. Les protéines selon l'invention sont donc susceptibles d'être impliquées dans la fonction de reproduction, mais également dans toutes les autres fonctions habituellement attribuées aux lipocalines.

10

Exemple 7 : Fabrication d'une protéine recombinante hOBPIIb α dans un système procaryote.

Une PCR sur l'ADN plasmidique d'un clone de hOBPIIb α à l'aide des amorces BIla/b (5' GTC GGA TCC CTG TCC TTC ACC CTG GAG G 3'), oligonucléotide sens démarquant 45 bases après l'ATG initiant la protéine, et XIIb (5' GTC CTC GAG GTG TTC GGG AAC GCA GCT TC 3'), oligonucléotide antisens précédant le codon stop de la protéine hOBPIIb α , a permis d'amplifier la totalité de l'ADN codant pour la protéine hOBPIIb α sécrétée. Les sites de restriction enzymatique BamH I et Xho I situés aux extrémités des deux oligonucléotides (bases soulignées), furent utilisés pour un clonage directionnel dans un vecteur plasmidique d'expression pGEX-6P1, suivi d'une transformation par électroporation (1800V, 200 Ω , 25 μ F) dans une souche bactérienne BL21. La synthèse de la protéine recombinante est obtenue par ajout d'IPTG à raison de 5 mM final pendant 3 h dans 250 ml de culture de la souche en milieu LB contenant 100 μ g/ml d'ampicilline préalablement incubée à 37°C pendant 2 h. Les cultures centrifugées sont reprises dans 25 ml de tampon TENGN. Le lysat est soniqué et centrifugé à nouveau. La protéine de fusion est alors purifiée à l'aide de 4 ml de billes fixant de façon covalente du glutathion insoluble (Sigma) pour 25 ml de surnageant. Après 4 h

d'incubation, elles sont lavées avec 3 volumes de NaCl 1M puis par 10 volumes de PBS 1X. L'élution est obtenue par mise en contact des billes pendant 10 min avec une solution de glutathion (Glutathion réduit 10 mM, Tris-HCl pH 8,0 50 mM).

5 La quantité de protéine recombinante produite est estimée par migration en gel de polyacrylamide et coloration au bleu de Coomassie (Figure 11). La spécificité de l'induction d'une protéine recombinante est testée par Western blot à l'aide d'un anticorps anti-GST de chèvre (Figure 11), révélé avec un anticorps anti-IgG de chèvre couplé à la 10 peroxydase en utilisant un kit ECL+plus (Amersham). La séparation de la protéine de fusion en deux protéines est obtenue par protéolyse de 100 µg de protéine recombinante préalablement dialysée à l'aide de 2U de preScission™ protéase (Pharmacia Biotech) dans 10 µl final contenant du tampon de l'enzyme 1X, pendant 4 h à 5°C.

15

Exemple 8 : Détection des hOBPII dans différents tissus par immunohistochimie.

La localisation des protéines hOBPII au sein des structures nasales (figure 12), des structures buccales et des glandes lachrymales 20 (figure 13), des glandes mammaires et des poumons (figure 14) est révélée par immunohistochimie.

25

REFERENCES

- Attwood, J. *et al.*(1994) Genomics 19, 203-214.
5
Barany, F., (1991), Proc. Natl. Acad. Sci. USA, 88, 189-193.
 Beste *et al.* (1999), Proc. Natl. Acad. Sci. USA, 96, 1898-1903
10 Bocskei, Z. *et al.*(1992) Nature 360, 186-188.

Borghoff S.J. *et al.* (1990) Annu. Rev. pharmacol. Toxicol. 30, 349-367

Burg, J.L. et al. (1996), Mol. and Cell. Probes, 10, 257-271.
15 Chu, B.C.F. et al.. (1986), Nucleic Acids Res., 14, 5591-5603.
Dear, T. N. *et al.* (1991) Biochemistry 30, 10376-10382.

20 Dewald, G. *et al.* (1996) Ann Hum Genet 60, 281-291.

Duck, P. *et al.* (1990), Biotechniques, 9, 142-147.

Edwards, C.P., and Aruffo, A., (1993), Curr. Op. Biotechnology 4, 558-
25 563.

Erlich, H.A., (1989), New York : Stockton Press.

Flower, D. R. (1996) Biochem J 318, 1-14.
30 Flower, D. R. (1995) J Mol Recognit 8, 185-195.

Gilbertson T.A. (1998) Cur. Op. Neurobiology 8, 447-452.

- Gregoire, C. *et al.* (1996) J Biol Chem 271, 32951-32959.
- 5 Guatelli J.C. *et al.* (1990), Proc. Natl. Acad. Sci. USA, 87, 1874-1878.
- Henzel, W. J. *et al.* (1988) J Biol Chem 263, 16682-16687.
- Holzfeind, P. and Redl, B. (1994) Gene 139, 177-183.
- 10 Hornigold, N. *et al.* (1997) Genomics 41, 385-389.
- Innis, M.A. et al. (1990), Academic Press.
- Igarashi, M. *et al.* (1992) Proc Natl Acad Sci U S A 89, 5376-2380.
- 15 Kievitis, T. et al. (1991), J. Virol. Methods, 35, 273-286.
- Kock, K. *et al.* (1994) Eur J Biochem 221, 905-916.
- 20 Kohler, G. et al.. (1975), Nature, 256 (5517), 495-497.
- Konieczny, A. *et al.* (1997) Immunology 92, 577-586.
- 25 Kwoh, D.Y. et al. (1989), Proc. Natl. Acad. Sci. USA, 86, 1173-1177.
- Lacazette, E. *et al.* (1997) Ann Hum Genet 61, 449-455.
- Landegren, U. *et al.* (1988), Science, 241, 1077-1080.
- 30 Lassagne, H. and Gachon, A. M. (1993) Exp Eye Res 56, 605-609.
- Lassagne, H. *et al.* (1995) Cytogenet Cell Genet 71, 104.

- Lassagne, H. *et al.* (1993) Genomics 18, 160-161.
- Lizardi, P.M. *et al.* (1988) Biotechnology 6, 1197-1202.
- 5 Luckow, V.A. (1993) Curr. Op. Biotechnology 4, 564-572.
- Matthews, J.A. *et al.* (1988) Anal. Biochem. 169 : 1-25.
- Miele, E.A. *et al.* (1983) J. Mol. Biol. 171 : 281-295.
- 10 Miller, W. L. (1998) Clin Perinatol 25, 799-817, v.
- Monaco, H. L. and Zanotti, G. (1992) Biopolymers 32, 457-465.
- 15 Nagata, A. *et al.* (1991) Proc Natl Acad Sci U S A 88, 4020-4024.
- Nahmias, J. *et al.* (1995) Eur J Hum Genet 3, 65-77.
- Pelosi, P. (1996) J Neurobiol 30, 3-19.
- 20 Pervaiz, S. and Brew, K. (1987) Faseb J 1, 209-214.
- Redl, B. *et al.* (1992) J Biol Chem 267, 20282-20287.
- 25 Rolfs, A., *et al.* (1991), Berlin : Springer-Verlag.
- Sambrook, J. *et al.* (1989), T. Sec. Ed. Cold Spring Harbor Lab., Cold Spring Harbor, New York.
- 30 Segev, D., (1992), Kessler C. Springer Verlag, Berlin, New-York, 197-205.
- Senoo, H. *et al.* (1990) J Lipid Res 31, 1229-1239.

- Shaw, P. H. *et al.* (1983) Cell 32, 755-761.
- Simard *et al.* (1992) Endocrinology 30, 1115-1121.
5
- Spinelli, S. *et al.* (1998) Biochemistry 37, 7913-7918.
- Stoesz S.P. and Gould M.N. (1995) Oncogene 11: 2233-2241.
- 10 Stone, B.B. *et al.* (1996) Mol. and Cell. Probes 10 : 359-370.
- van Slegtenhorst, M. *et al.* (1995) Eur J Hum Genet 3, 78-86.
- 15 Walker, G.T. *et al.* (1992) Nucleic Acids Res. 20 : 1691-1696.
- Zeng, C. *et al.* (1996) Proc Natl Acad Sci U S A 93, 6626-6630.

LISTAGE DE SEQUENCES

<110> UNIVERSITE D'AUVERGNE CLERMONT I
 5 <120> « ODORANT-BINDING » PROTEINES HUMAINES FIXANTS DES LIGANDS
 HYDROPHOBES :
 POLYPEPTIDES ET POLYNUCLEOTIDES CODANT LESDITS POLYPEPTIDES, ET
 LEURS APPLICATIONS
 10
 <130> D15947
 15 <160> 16
 <170> PatentIn Vers. 2.0
 20 <210> 1
 <211> 676
 <212> ADN
 <213> Homo sapiens
 25 <220>
 <221> CDS
 <222> (43)..(552)
 <223> cDNA396 (676) /g1 (hOBPIIa-alpha)
 30 <400> 1
 cggccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg
 Met Lys Thr Leu
 1
 ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg
 Phe Leu Gly Val Thr Leu Ala Ala Leu Ser Phe Thr Leu
 35 5 10 15 20
 102 ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg
 Phe Leu Gly Val Thr Leu Ala Ala Leu Ser Phe Thr Leu
 5 10 15 20
 gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc
 Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
 25 30 35
 150 gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg
 Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
 40 45 50
 198 gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg
 Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val
 45 50
 246 aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc
 Lys Val Thr Ala Leu Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe
 45 55 60 65
 294 atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg
 Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr
 50 55 75 80
 342 gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac
 Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Arg Lys Leu Ile Tyr
 55 85 90 95 100
 39 ctg cag gag ctg ccc ggg acg gac gac tac gtc ttt tac tgc aaa gac

Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp
 105 110 115

5 cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggt agg aat 438
 Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Arg Asn
 120 125 130

10 cct aat acc aac ctg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag 486
 Pro Asn Thr Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln
 135 140 145

15 cac aag gga ctc tcg gag gag gac att ttc atg ccc ctg cag acg gga 534
 His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly
 150 155 160

20 agc tgc gtt ctc gaa cac taggcagccc ccgggtctgc acctccagag 582
 Ser Cys Val Leu Glu His
 165 170

25 cccaccctac caccagacac agagccccga ccacacctggac ctaccctcca gccatgaccc 642
 ttccctgctc ccacccacct gactccaaat aaag 676

30 <210> 2
 <211> 170
 <212> PRT
 <213> Homo sapiens

35 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15

40 Ser Phe Thr Leu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
 20 25 30

45 Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
 35 40 45

50 Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Asn Leu Glu Ala
 50 55 60

55 Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65 70 75 80

60 Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
 85 90 95

65 Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
 100 105 110

70 Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
 115 120 125

75 80 85 90 95

80 85 90 95

85 90 95

90 95

95

100 105 110

115 120 125

130 135 140

140

Lys Leu Val Gln His Lys Gly Leu Ser Glu Glu Asp Ile Phe Met Pro

	145	150	155	160	
	Leu Gln Thr Gly Ser Cys Val Leu Glu His				
		165	170		
5					
	<210> 3				
	<211> 725				
	<212> ADN				
10	<213> Homo sapiens				
	<220>				
	<221> CDS				
	<222> (43)..(480)				
15	<223> cDNA396 (725) /SM12 (hOBPIIa-beta)				
	<400> 3				54
	cgtccaggta cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg				
		Met Lys Thr Leu			
		1			
20	ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg				102
	Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu				
		5	10	15	20
25	gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc				150
	Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val				
		25	30	35	
30	gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg				198
	Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val				
		40	45	50	
35	aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc				246
	Lys Val Thr Ala Leu Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe				
		55	60	65	
40	atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg				294
	Met Arg Glu Asp Arg Cys Ile Gln Lys Ile Leu Met Arg Lys Thr				
		70	75	80	
45	gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac				342
	Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg Lys Leu Ile Tyr				
		85	90	95	100
50	ctg cag gag ctg ccc ggg acg gac tac gtc ttt tac tgc aaa gac				390
	Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp				
		105	110	115	
55	cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg ggg cgc tgc				438
	Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Gly Pro Cys				
		120	125	130	
	cgc tgt ccc cac gtc ggc tca cct ggc cac ctc acc tgc agg				480
	Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr Cys Arg				
		135	140	145	
	taggaatcct aataccaacc tggaggccct ggaagaattt aagaaattgg tgcatcacaa				540

gggactctcg gaggaggaca ttttcatgcc cctgcagacg ggaagctgcg ttctcgaaca 600
 5 ctaggcagcc cccgggtctg cacctccaga gcccacccta ccaccagaca cagagccccgg 660
 accacacctgga cctaccctcc agccatgacc cttccctgct cccacccacc tgactccaaa 720
 taaag
 10 725
 <210> 4
 <211> 146
 <212> PRT
 <213> Homo sapiens
 15
 <400> 4
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15
 20 Ser Phe Thr Leu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
 20 25 30
 Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
 35 40 45
 25 Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Asn Leu Glu Ala
 50 55 60
 30 Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65 70 75 80
 Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
 85 90 95
 35 Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
 100 105 110
 Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
 115 120 125
 40 Val Gly Pro Cys Arg Cys Pro His Val Gly Ser Pro Gly His Leu Thr
 130 135 140
 Cys Arg
 45 145
 <210> 5
 <211> 741
 50 <212> ADN
 <213> Homo sapiens
 <220>
 <221> CDS
 55 <222> (43)..(726)
 <223> cDNA396 (741) /SM4 (hOBPIIa-gamma)
 <400> 5

	cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg Met Lys Thr Leu 1	54
5	tgc ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Leu Ser Phe Thr Leu 5 10 15 20	102
10	gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val 25 30 35	150
15	gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val 40 45 50	198
20	aag gtg aca gcc ctg ggc ggt ggg aac ttg gaa gcc acg ttc acc ttc Lys Val Thr Ala Leu Gly Gly Asn Leu Glu Ala Thr Phe Thr Phe 55 60 65	246
25	atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr 70 75 80	294
30	gag gag cct ggc aaa ttc agc gcc tat ggg ggc agg aag ctc ata tac Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Arg Lys Leu Ile Tyr 85 90 95 100	342
35	ctg cag gag ctg ccc ggg acg gac tac gtc ttt tac tgc aaa gac Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe Tyr Cys Lys Asp 105 110 115	390
40	cag cgc cgt ggg ggc ctg cgc tac atg gga aag ctt gtg gca tct gct Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu Val Ala Ser Ala 120 125 130	438
45	ccc tgc agg gcc gtg ccg ctg tcc cca cgt cgg ctc acc tgg cca cct Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu Thr Trp Pro Pro 135 140 145	486
50	cac ctg cag gta gga atc cta ata cca acc tgg agg ccc tgg aag aat His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg Pro Trp Lys Asn 150 155 160	534
55	tta aga aat tgg tgc agc aca agg gac tct cgg agg agg aca ttt tca Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg Arg Thr Phe Ser 165 170 175 180	582
	tgc ccc tgc aga cgg gaa gct gcg ttc tcg aac act agg cag ccc ccg Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr Arg Gln Pro Pro 185 190 195	630
	ggt ctg cac ctc cag agc cca ccc tac cac cag aca cag agc ccg gac Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr Gln Ser Pro Asp 200 205 210	678
	cac ctg gac cta ccc tcc agc cat gac cct tcc ctg ctc cca ccc acc His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu Leu Pro Pro Thr	726

215

220

225

tgactccaaa taaag 741

5 <210> 6
 <211> 228
 <212> PRT
 <213> Homo sapiens

10 <400> 6
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15

15 Ser Phe Thr Leu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys
 20 25 30

Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys
 35 40 45

20 Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Asn Leu Glu Ala
 50 55 60

25 Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu
 65 70 75 80

Met Arg Lys Thr Glu Glu Pro Gly Lys Phe Ser Ala Tyr Gly Gly Arg
 85 90 95

30 Lys Leu Ile Tyr Leu Gln Glu Leu Pro Gly Thr Asp Asp Tyr Val Phe
 100 105 110

Tyr Cys Lys Asp Gln Arg Arg Gly Gly Leu Arg Tyr Met Gly Lys Leu
 115 120 125

35 Val Ala Ser Ala Pro Cys Arg Ala Val Pro Leu Ser Pro Arg Arg Leu
 130 135 140

40 Thr Trp Pro Pro His Leu Gln Val Gly Ile Leu Ile Pro Thr Trp Arg
 145 150 155 160

Pro Trp Lys Asn Leu Arg Asn Trp Cys Ser Thr Arg Asp Ser Arg Arg
 165 170 175

45 Arg Thr Phe Ser Cys Pro Cys Arg Arg Glu Ala Ala Phe Ser Asn Thr
 180 185 190

Arg Gln Pro Pro Gly Leu His Leu Gln Ser Pro Pro Tyr His Gln Thr
 195 200 205

50 Gln Ser Pro Asp His Leu Asp Leu Pro Ser Ser His Asp Pro Ser Leu
 210 215 220

55 Leu Pro Pro Thr
 225

<211> 607
 <212> ADN
 <213> Homo sapiens
 5 <220>
 <221> CDS
 <222> (43)..(483)
 <223> cDNA396 (607) - forme courte (hOBPIIa-delta)
 10 <400> 7
 cgccccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg Met Lys Thr Leu 54
 15 ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu 102
 20 5 10 15 20
 gag gag gag gat gag gga gaa tcg gtg cat cca gaa gaa aat cct gat 150
 Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu Glu Asn Pro Asp
 25 25 30 35
 25 gcg gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa 198
 Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu
 40 40 45 50
 30 gct cat ata cct gca gga gct gcc cgg gac cga cta cgt ctt tta 246
 Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg Leu Arg Leu Leu
 55 55 60 65
 35 ctg caa aga cca gcg ccg tgg ggg cct gcg cta cat ggg aaa gct tgt 294
 Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His Gly Lys Ala Cys
 70 70 75 80
 40 ggc atc tgc tcc ctg cag ggc cgt gcc gct gtc ccc acc ttg gct cac 342
 Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro Thr Leu Ala His
 85 85 90 95 100
 45 ctg gcc acc tca cct gca ggt agg aat cct aat acc aac ctg gag gcc 390
 Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr Asn Leu Glu Ala
 105 105 110 115
 50 ctg gaa gaa ttt aag aaa ttg gtg cag cgc aag gga ctc tcg gag gag 438
 Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu
 120 120 125 130
 55 gac att ttc atg ccc ctg cag acg gga agc tgc gtt ctc gaa cac 483
 Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val Leu Glu His
 135 135 140 145
 55 taggcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgga 543
 ccacccctggac ctaccctcca gccatgaccc ttccctgctc ccacccacct gactccaaat 603
 aaag 607

<211> 147
 <212> PRT
 <213> Homo sapiens

5 <400> 8
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15

10 Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Glu Ser Val His Pro Glu
 20 25 30

Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
 35 40 45

15 Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Arg Asp Gly Arg
 50 55 60

Leu Arg Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Leu His
 65 70 75 80

20 Gly Lys Ala Cys Gly Ile Cys Ser Leu Gln Gly Arg Ala Ala Val Pro
 85 90 95

25 Thr Leu Ala His Leu Ala Thr Ser Pro Ala Gly Arg Asn Pro Asn Thr
 100 105 110

Asn Leu Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln Arg Lys Gly
 115 120 125

30 Leu Ser Glu Glu Asp Ile Phe Met Pro Leu Gln Thr Gly Ser Cys Val
 130 135 140

Leu Glu His
 145

35

<210> 9
 <211> 676
 <212> ADN
 40 <213> Homo sapiens

<220>
 <221> CDS
 <222> (43)..(552)
 45 <223> cDNA2098 (676) - forme classique (hOBPIIb-alpha)

<400> 9
 cgcccaagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg
 50 Met Lys Thr Leu
 1

ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg
 Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu
 5 10 15 20 102

55 gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg gtg gtc
 Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met Val Val
 25 30 35 150

	gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc cca gtg Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser Pro Val 40 45 50	198
5	aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc acc ttc Lys Val Thr Ala Leu Gly Gly Lys Leu Glu Ala Thr Phe Thr Phe 55 60 65	246
10	atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg aag acg Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg Lys Thr 70 75 80	294
15	gag gag cct ggc aaa tac agc gcc tat ggg ggc agg aag ctc atg tac Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg Lys Leu Met Tyr 85 90 95 100	342
20	ctg cag gag ctg ccc agg agg gac cac tac atc ttt tac tgc aaa gac Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe Tyr Cys Lys Asp 105 110 115	390
	cag cac cat ggg ggc ctg ctc cac atg gga aag ctt gtg ggt agg aat Gln His His Gly Gly Leu Leu His Met Gly Lys Leu Val Gly Arg Asn 120 125 130	438
25	tct gat acc aac cgg gag gcc ctg gaa gaa ttt aag aaa ttg gtg cag Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys Lys Leu Val Gln 135 140 145	486
30	cgc aag gga ctc tcg gag gag gac att ttc acg ccc ctg cag acg gga Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro Leu Gln Thr Gly 150 155 160	534
35	agc tgc gtt ccc gaa cac taggcagccc ccgggtctgc acctccagag Ser Cys Val Pro Glu His 165 170	582
	ccccaccctac caccagacac agagcccgga ccacctggac ctaccctcca gccatgaccc	642
40	ttccctgctc ccacccacct gactccaaat aaag	676
	<210> 10	
	<211> 170	
45	<212> PRT	
	<213> Homo sapiens	
	<400> 10	
	Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu	
50	1 5 10 15	
	Ser Phe Thr Leu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys	
	20 25 30	
55	Ala Met Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys 35 40 45	
	Val Ser Pro Val Lys Val Thr Ala Leu Gly Gly Lys Leu Glu Ala	

	50	55	60	
5	Thr Phe Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu 65 70 75 80			
	Met Arg Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Tyr Gly Gly Arg 85 90 95			
10	Lys Leu Met Tyr Leu Gln Glu Leu Pro Arg Arg Asp His Tyr Ile Phe 100 105 110			
	Tyr Cys Lys Asp Gln His His Gly Gly Leu Leu His Met Gly Lys Leu 115 120 125			
15	Val Gly Arg Asn Ser Asp Thr Asn Arg Glu Ala Leu Glu Glu Phe Lys 130 135 140			
20	Lys Leu Val Gln Arg Lys Gly Leu Ser Glu Glu Asp Ile Phe Thr Pro 145 150 155 160			
	Leu Gln Thr Gly Ser Cys Val Pro Glu His 165 170			
25	<210> 11 <211> 782 <212> ADN <213> Homo sapiens			
30	<220> <221> CDS <222> (1)..(537) <223> cDNA2098 (782) - forme longue (hOBPIIb-beta)			
35	<400> 11 cgc cca gtg acc tgc cga ggt cgg cag cac aga gct ctg gag atg aag Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys 1 5 10 15			48
40	acc ctg ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe 20 25 30			96
45	acc ctg gag gag gag gat atc aca ggg acc tgg tac gtg aag gcc atg Thr Leu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met 35 40 45			144
50	gtg gtc gat aag gac ttt ccg gag gac agg agg ccc agg aag gtg tcc Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser 50 55 60			192
	cca gtg aag gtg aca gcc ctg ggc ggt ggg aag ttg gaa gcc acg ttc Pro Val Lys Val Thr Ala Leu Gly Gly Lys Leu Glu Ala Thr Phe 65 70 75 80			240
55	acc ttc atg agg gag gat cgg tgc atc cag aag aaa atc ctg atg cgg Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg 85 90 95			288

aag acg gag gag cct ggc aaa tac agc gcc tgc ttg tcc gca gtc gag Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu 100 105 110	336
5 atg gac cag atc acg cct gcc ctc tgg gag gcc cta gcc att gac aca Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr 115 120 125	384
10 ttg agg aag ctg agg att ggg aca agg agg cca agg att aga tgg ggg Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly 130 135 140	432
15 cag gaa gct cat gta cct gca gga gct gcc cag gag gga cca cta cat Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His 145 150 155 160	480
20 ctt tta ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys 165 170 175	528
25 gct tgt ggg taggaattct gataccaacc gggaggccct ggaagaattt Ala Cys Gly	577
30 aagaaaattgg tgcagcgcaa gggactctcg gaggaggaca tttcacgcc cctgcagacg 637 ggaagctgcg ttcccgaaca ctaggcagcc cccgggtctg cacctccaga gcccacccta 697 ccaccagaca cagagccccgg accacacctgga cctaccctcc agccatgacc cttccctgct 757 cccacccacc tgactccaaa taaag	782
35 <210> 12 <211> 179 <212> PRT <213> Homo sapiens	
40 <400> 12 Arg Pro Val Thr Cys Arg Gly Arg Gln His Arg Ala Leu Glu Met Lys 1 5 10 15	
45 Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe 20 25 30	
50 Thr Leu Glu Glu Glu Asp Ile Thr Gly Thr Trp Tyr Val Lys Ala Met 35 40 45	
55 Val Val Asp Lys Asp Phe Pro Glu Asp Arg Arg Pro Arg Lys Val Ser 50 55 60	
60 Pro Val Lys Val Thr Ala Leu Gly Gly Lys Leu Glu Ala Thr Phe 65 70 75 80	
70 Thr Phe Met Arg Glu Asp Arg Cys Ile Gln Lys Lys Ile Leu Met Arg 85 90 95	
75 Lys Thr Glu Glu Pro Gly Lys Tyr Ser Ala Cys Leu Ser Ala Val Glu	

	100	105	110
5	Met Asp Gln Ile Thr Pro Ala Leu Trp Glu Ala Leu Ala Ile Asp Thr 115	120	125
	Leu Arg Lys Leu Arg Ile Gly Thr Arg Arg Pro Arg Ile Arg Trp Gly 130	135	140
10	Gln Glu Ala His Val Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His 145	150	155
	Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys 165	170	175
15	Ala Cys Gly		
20	<210> 13 <211> 542 <212> ADN <213> Homo sapiens		
25	<220> <221> CDS <222> (43)..(297) <223> cDNA2098 (542) - forme courte (hOBPIIb-gamma)		
30	<400> 13 cgcccagtga cctgccgagg tcggcagcac agagctctgg ag atg aag acc ctg Met Lys Thr Leu		54
	1		
35	ttc ctg ggt gtc acg ctc ggc ctg gcc gct gcc ctg tcc ttc acc ctg Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu Ser Phe Thr Leu 5 10 15 20		102
40	gag gag gag gat gag gga gga tcg gtg cat cca gaa gaa aat cct gat Glu Glu Glu Asp Glu Gly Ser Val His Pro Glu Glu Asn Pro Asp 25 30 35		150
45	gct gaa gac gga gga gcc tgg caa att cag cgc cta tgg ggg cag gaa Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu Trp Gly Gln Glu 40 45 50		198
50	gct cat ata cct gca gga gct gcc cag gag gga cca cta cat ctt tta Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro Leu His Leu Leu 55 60 65		246
55	ctg caa aga cca gca cca tgg ggg cct gct cca cat ggg aaa gct tgt Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His Gly Lys Ala Cys 70 75 80		294
	ggg taggaattct gataccaacc gggaggccct ggaagaattt aagaaaattgg Gly 85		347
	tgcagcgcaa gggactctcg gaggaggaca tttcacgcc cctgcagacg ggaagctgcg	407	

ttcccgaaca cttaggcagcc cccgggtctg cacctccaga gcccacccta ccaccagaca 467
 cagagcccg accacacctgga cctaccctcc agccatgacc cttccctgct cccacccacc 527
 5 tgactccaaa taaag 542

10 <210> 14
 <211> 85
 <212> PRT
 <213> Homo sapiens

15 <400> 14
 Met Lys Thr Leu Phe Leu Gly Val Thr Leu Gly Leu Ala Ala Ala Leu
 1 5 10 15

Ser Phe Thr Leu Glu Glu Glu Asp Glu Gly Gly Ser Val His Pro Glu
 20 25 30

20 Glu Asn Pro Asp Ala Glu Asp Gly Gly Ala Trp Gln Ile Gln Arg Leu
 35 40 45

25 Trp Gly Gln Glu Ala His Ile Pro Ala Gly Ala Ala Gln Glu Gly Pro
 50 55 60

Leu His Leu Leu Leu Gln Arg Pro Ala Pro Trp Gly Pro Ala Pro His
 65 70 75 80

30 Gly Lys Ala Cys Gly
 85

35 <210> 15
 <211> 10664
 <212> ADN
 <213> Homo sapiens

40 <220>
 <223> Gène hOBPIIa

<400> 15
 ctctgcttct atattcggcc tccaaatctg tgctcaatgc agcaactgga gtgaccctt 60

45 aaatacgtaa gtcacagctt gcctttgtca gagctatcca gggctttca ctcagagcag 120
 aagctgaagt cctcgtggtg gtccttaatc cctacatggc cggtccaccc actccccagc 180

50 ctcatgtgtg gccggctctcc ctggatcatt tgctgtggct gctctgccgt gtgttcccg 240
 aaactgccag cgttctccca cctctggct ggcactggat gctcctgcca cctggactcc 300

tcttccagct gacgagctca tggcttgctt cttcatgtc ttaaattcgg tgtttgaatg 360

55 ccaccttggc gaggctgttc ctcatcaatt catgtaaagg acaaataacc ctttgtctgg 420
 ccactcctgt ccaacttgc cactttgctt tttccatagc actgatcacc atttaaaata 480

atgttatgaac cagtgaccca tcagtgacca gagacagtag attgaacagg tattgcttgc 540
acagtgaaca atcagctcta gaggttaaa gagaaccaca aagaatggct caggattgca 600
5 cgagggcagt taaggaagaa ataaacaggg gtggggatcc ttccacaag tgtgggttca 660
gaccgcattgg gagaagggtgt gggtctccca agggaaagctg gagaagtttgc 720
10 ccagccacag ctggcccacg gtcagggcag aggccagcag gcagggaaat gtaggctggg 780
tctggcaagc aggaagcctc tcttcccccg caaggcaggt gggctgggc tggaaagcct 840
acaagagtca ctgggaaccc acaggtgcag atgccactga atctcaatag gaagccatct 900
15 gggggggtcc cctgaattca atgtggtgtg tgccgccaga cgtcccaac ttgtgccact 960
gccatttata caggaagaga aagaaaaagg aagaaatgga agcatctgga accagtcatc
1020
20 ctggaccat gctaggaggt gcttccccc acgcctcaac caacaaagct tagcattgca
1080
ccaggtgcac aagagaagcg cctgctggct ccagctccat tgccctcgaa ccagccatga
1140
25 agggtgctgt tggagctgga ggcaagacat tgatagctgg cactgcaatt cacttattta
1200
30 ttttgttcat tttaagtccc ctgcacccat aatataagcc ccccaagcac aggacatttgc
1260
ttttatttgcgtt cgtatgtattt cttgtcccc aaagaatgag aggcatctgaa aagatctgca
1320
35 aaaatcaaac ataaaaatga acctttattt cgtatgtttt attttgcgtt atatgg
1380
catttccaaat attttgcgtt aacaatttgcgtt cccttctccgtt gttcatgtg tggagtt
1440
40 tctgttcaaa tgatgttttctt gaggatggatt gctgaggctt tggctctagg tttttttttt
1500
45 ttaagcattt atgctcattt tggttttttttaa attaaacattt taaccctgag acactgtaga
1560
ttcccatgca attgtaaaaa gccacacaga gctatcatgt gtatcttccac ctggcttgc
1620
50 ccagccccaa ccccgaaat gacagacctg ttctccactc ctgcaatctg ctcattgca
1680
gaatgtcgatc tatttgcattt ataaaaattgtt gggattggct ttttttttttcc tttttttttt
1740
55 catttcttgcgtt agatttcatcc tatttgcattt tttatcaata gtttatttccca ttttacttct
1800

gagtagtgct ctatggtagt gatgtaccac agtctgttta accattcacc tgttggagga
1860
5 agtctgttta tatacgatgg ggctatgaca catgtatagg ttttgcattt gacatcgatgg
1920
10 ttcatttccc tgggacaaag gcccagggggt tctattgctg gattctatgc ttgttacagg
1980
15 gctcattttg ttttgttttgc ctttgtttta ttttgttttta aacctgtcaa gccatttcc
2040
2100 agatccagggtt tccatgcattt ctcaccaggc ttcaggatgtatgc tcactatgtat cttatctcagg
2160 ccacctaataat aaggatgtac tgatatatca tggcttttat ttgcatttca ctgatgacta
2220 atggtgttga gcatttttc atgtgttttat ttgcattctg tatacctctt ctatgtcaatgt
2280 :
gtcccttcattt gtctttgtt tacgttctat ttttggaaact gttgagtttt gaaaaattct
2280
2340 ttataaagta tagaaactaa ttctttgtt aatatgtatgt ttgtcaatata tttctttcagg
2400 :
tctttggctt gtcttttat tctgttaaca gggctcttta cagagcaaaa gttttttatt
2460 :
ttgatgaagt ctattttaac aatttttcct ttatggatc atattttgg caacaaatct
2520 aaaacccctt gacccagctc catatgtcaa agatttctt gttttctaaa agttttatag
2580 :
tttaagttt tatgttgaag tctatgatcc attttgagtt aattttcata cagggtgtga
2640 gaccaagggtt gtggttcttc ttttcttttgc gtgggtttgt ctgtggatgt ccagttgtc
2700 :
cagcccatttt gctacaaaag ctatcttcc tccactgaat tacttttgc tctttgtaaa
2760 :
aatgttaattt ggtgtatgg tacaggatgt tttgaggattt ctttattgtc tgtccccatttgc
2820 :
atctatgcat ctgtccatgt gctagctata taagtcttgc aagggttagcc ttttagctgg
2880 :
tgtgggtggcg agagcctgtt aatcccagctt ttcggggaaatggggcaggaa gaatcgatttgc
2940 :
aacccaggag gtggaggatgttgc cagtggatgtca agatcggtcc actgcactcc agcctgggttgc

acagagtcaa ctccatataa aaaaataaaa acaaaaataaa agtagcgtga 3000	:tctcccac
5 cttattcttt tttaaaaaag ttttagctat tccagttcct ttgcctttcc 3060	:ataaaattt
tagaataatc ttgtctatat ctaccacaat tctttctgga attttgatag 3120	:aattttgtt
0 acatcttat attatggag agaaatgata tttttactat gttgagtctt 3180	:aatccatg
5 gatataacgt ctctccattt gtttagatct tctttgattt attttataat 3240	:ttgcattg
ttttcagcat acaaattccag catatgggtt gtttagactt tgccctaggca 3300	:tcattttt
0 ttagccatta taaatagtac tgtgtttta agtttagggg ccattactag 3360	:atgtagaca
cacaattgtat ctttgtatcat ttatcttgc tcctgcaacc ttgctgaact 3420	:acttactag
5 ttcttaggagg tgtattgttt tctttgttt tgttttcaa tttcttggga 3480	:tttctacag
) agataataat gtcatctgca gatgcagttt tcttccttcc tttccaattt 3540	:tatgtcttt
aatttccttt tttaaaaacc tatattactc tgactagaac tttctgtact 3600	:tgttaaata
; caagtggtag agtggacatc cttgccttgt ccctgatgtt aaagagaaaag 3660	:atttgtaac
) tgagtatccc agcctcaaaa tgtgctaaa aactttttc ctttcttgct 3720	:tcagccttg
aaacataactt cgaaactctt tatttctccc tttccacca ggcacttccg 3780	:gagcagtgc
tcgcttatct aattatgtgc ttacttagaa attccagggg ccaattttga 3840	:acaaaccag
gcagagagac ccagctgcag aatcctccct ctttaggggaa gttacaggtt 3900	:cctaccact
tccccggctga aatcaggatg acgcaaacc aacccctccgga cagacgattt 3960	:tgactcaca
ataaccatca gaacaagatg cagaccaaca tcctcctgca ccattccac 4020	:tatttccca
cacctttcc tccttaaacc cttcgctca gtccagaaaa tctgaatggt 4080	:ttttaaagg

catgggtctg gccattcccc aactgccagt atttgaataa agctgcttc ctttaccac
4140

5 acctcacttc tcatgccttg acttctgagc agcgagcagc tggacttgag ccagttacac
4200

atcagtcttt caccattaga aataatgtag ccataggttt tttttcgta gatgttctt
4260

10 ttcaagttaa agaagttctc ttctattcct attttctga gaggtgttat cccgaatgag
4320

tgttgaattt tgttaaatac ttttaacaac ccaacaggaa ccaccatcag aagccatcca
4380

15 gagaaacgca ccaggccaca aagcactggg ggccagggat ctgcctctg ctgtctgcca
4440

20 tgggttacc cccagcctcc aaccctacca tcccctgacg gtgtctgcag cagttgaacc
4500

caaccagcat ctaaaagaac acagtttgt aacgagactg ggacacaggg caagatgggt
4560

25 ggacaatggg aggctcctgg agagcacccg taccagcgag catagaattc atgggggtga
4620

cctgttccct gaagcatctg cgctgttgt tccagcattt tcttcaagga ttgagccagc
4680

30 agcaccagtg tcatacggtg cttaaatcaa tgattcacag ccaaccaatg aaatacaagg
4740

tgccggctgg gcgcggtgtgc tcacgcctgt aatcccagca ctttggagg ccgaggcggg
4800

35 cgatcacga ggtcaggaga tcgagaccat cccggctaaa acggtaaac cccgtctcta
4860

40 ctaaaaatac aaaaaaaatt agccggcgt agtggcgggc gcctgttagtc ccagctactt
4920

gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg gagcttgcag tgagccaaga
4980

45 tcccgccact gcactccagc ctggcgaca gagcgagact ccgtctcaaa aaaaaaaaaa
5040

aaaaaaaaaaa aaaagaaata caaggtgcct ggggtacagg caacaaaaat ggggaacgga
5100

55 gaaatcttcc cggtgaggta gagactccag cctgtttgct catctctggt cctgaagagc
5160

ccagtccccg ccctaaggat ggggtttctg ctcggcagca cttgccgtga gaggggtgag
5220

gcactgggtc acgccagccc ttctttata gccccaggg tttatttagag tgtgcattag
 5280

5 tattattnag caagcacttg agggtgtctg atcttggcc agagaggcac aaagattgtg
 5340

tgggcccagc acctgcctgc agagcgtggg tcagccttgg gtcccaggc agatgacgca
 5400

10 ggccccggaa ctacagggtc caggagggag aaggcaaagt tctggctcag gttggctgg
 5460

gatgaggcca gcggagccag gtgccaagg gagctcagcc acaaactctg agcacaggct
 5520

15 ggcaggtggc tcttgatgct catgccaccc atttatctaa agggatgaga ttcaaggcct
 5580

20 gtcctgggtgg ctggggcccg tcgaagctga cgagagaggg gatgttagagt gaatgtatat
 5640

tccactctac cacttgtatt taacagggag atggatgatg aacacgtgca ggaggaaaca
 5700

25 ggcaggacaa tccagagaga tcacgtgttc tgaggacagc acagccaggc tccggtaacgg
 5760

30 agtgaagcgg ggtggggcag gcggcggggt ccctcatatg gcccggaggag gccgtatata
 5820

tactgacctt gagccacaca atagtgcct tctctgcccc taggaagctc gagtgcagg
 5880

35 tccaggtggg gaaaatcaat gcagagtggg tcccagagtg ggccggaaagct tgggctctag
 5940

ggcgtgcggg actcagctgg cagcagcccc acatctctat gttctaaag cccagtcac
 6000

40 ttctgctcag cagggagatg gccagttccc cagaggactt gcccagggtc ccagccgtgg
 6060

ccctgggagg gctcagggtt gagggaggga gaattccgag ccgtggccccc tccctctgtg
 6120

45 gctcgggtg cagacgtgca ctaccctgcc ctgtcctctg gagtcctccg ggctcatcca
 6180

50 agtgggcgca tttgggtgca aaccctggac agcatcgtgt gtctttctt ctctggctgg
 6240

cactgaattt ctgttacca gatgccagca atgaggacac tccccctccc ggatgcagga
 6300

55 ctgggtgcca ccatggggag gggtgcaccc tgccacgtgc tcccaggaca ctggccagg
 6360

ggctataaaag aacatctcg aaggagccag cacagccttg ttcagacgcc cagtgacctg
6420

5 ccgagggtcg cagcacagag ctctggagat gaagaccctg ttccctgggtg tcacgctcg
6480

cctggccgct gccctgtcct tcaccctgga ggaggaggat gtgagctggg ttggcgtggg
6540

10 cgatggagg agccaggtgg actcctgggc agggggcagt gccaggggcc ctgctttagg
6600

aggtgtcaact taagcctggg gtgggtggtgg aggggtccta ttgttccttc agcagacaat
6660

15 gctccccatg aggcccagg ccggagcagg ctcggctcag gggttcctgc tgcactgacg
6720

cctgaagccc gaagggtctcg cagggttggg ccctgtggag ggagggctca cctggtgctg
20 6780

ggcccccgggg gtcacatgggg tgcagacatg ccctccttcc actggggct gggagccctg
6840

25 agcagggggc tggctctaac tcactccagc tgagctctaa ctaaggtgca ggaacccagc
6900

ctgcctttag gggtggcagc cgggcaccat gggtgtctgg ttatagctgc aggctgagt
6960

30 gccagggtca gagtagaatac tgggccaccc atggtgggct cacggccttg gcctgctcca
7020

gatcacaggg acctggtacg tgaaggccat ggtggtcgat aaggacttgc cgaggacag
35 7080

gaggcccagg aaggtgtccc cagtgaaggt gacagccctg ggcggtggga acttgaaagc
7140

40 cacgttccacc ttcatgtgag tggccac tgcagggccc ctcaggccac ttgcctccc
7200

cgcggccagac ccacctggtg cccattgccc catccacatt tcgggtgttg ggaagagtca
7260

45 ccccccgcct tggagggaaa cagccagggc atcctgaagc tcgggtgggt gggggggcag
7320

50 tggaaatttc agttgccgg gtcagggcca tgcaccaggt gagctgagga tggccaggt
7380

gtgtcctggg agccgctgcc cgcgtgtctc ctgtttcca ggagggagga tcggtgcatc
7440

55 cagaagaaaa tcctgatgctg gaagacggag gacgcctggca aatacagcgc ctgtgagccc
7500

ctccccgacc ccccactccc catgcccaac cccggatgca ccagccccac tgcagggtgga
 7560

5 gagtgcccaag gccacacttc tgccagggtc ccagccctgc ccacctccaa ggaggggctg
 7620

gcctctcctt cctggggggc tggtgccct gacatcagac accaggtgtg acaggcttgt
 7680

10 ccacagtaga gatggaccag atcaaggctg ccctctgggaa ggccctagcc attgacacat
 7740

tgaggaatcc gagtgtagg gaccaggagg ccgagggtta gggatggaa gccaaggctg
 7800

15 agggtttggg atcaggaatc cgagggttag ggacaggaa gtggggcagg agcagctgt
 7860

20 ggagctgggaa aggccggact ctagtcctgg acgtgctctg gccttgtggc tccattactt
 7920

gcattgggac cttccgagag gaggctcctg cctccgtgtc cgggtccatg ctgtgcggag
 7980

25 cagccaggcc tggctcaggc tgtccaggc acctgggtga ccactgaaac attcctgagt
 8040

30 gtttcttcgt gtggcctga gtgctctctc cggaaatgag ggcactgaag acccatcttc
 8100

tctgtcatct acagatgggg gcaggaagct catatacctg caggagctgc ccggacgga
 8160

35 cgactacgtc ttttactgca aagaccagcg ccgtggggc ctgcgtaca tggaaagct
 8220

tgtgggtgag gggcccgctg gggcctgcat gtcctgcccc atggtctctg cctccagaag
 8280

40 ccagtggAAC caccatcatc acgcctggc acggggggaa aaggaagccc cctgcgcgg
 8340

ccttcgtgtc ctaggcacca agcgctgccc tggatggctg gtccaaatgc ctgaagtggg
 8400

45 agtgggggtgg gccaggcagg gacagacacg gcccctcggtg acgtgaacct gccaaggccc
 8460

50 gcttgtgggg tctcaggtgt aggggcctca ctttaagggg gaggtancat cttAACAGAG
 8520

ctttcatgg ggcaggact ctccaggcg gcaggcagc cagtcctct gggacacaag
 8580

55 gtcctccag gtgagggttg tgaccctgca gagtggctt gggagctgcc caggtcccc
 8640

tggggttgct gagtggcttg gaccctgccca ctgtcccctt tcctggggac ctctcacctg
8700

5 ggcggtgcc gcctcctctg tccccagtc caccctgag ctcttgtcca ttctcaggcc
8760

 tcctctcccc cttgcctggt gctggacagt tgccatctct tctgtccccca gccccacccc
8820

10 tgagctctga tccactctcg ggcctctccc ccgtcctgat gctggccgt ggtcgctcc
8880

 tttagcatc tgctccctgc agggccgtgc cgctgtcccc acgtcggctc acctggccac
8940

15 ctcacctgca ggttaggaatc ctaataccaa cctggaggcc ctggaagaat ttaagaatt
9000

20 ggtcagcac aagggactct cggaggagga cattttcacg cccctgcaga cgggtgagga
9060

 cggctgtgcc cagtaccccg tttcccttg tgtctctgtg tcatctccag tgtcccatga
9120

25 ccctcgtgtc ctcccatgtc cccgcattc cccatgtgcc ccgagtcctcc tcgcaggggc
9180

 tctggccct gcttagcatc ctgcgtttg gagggtctgc actctggct gcgatggggt
9240

30 ctggggctcc gcgtctggg ctgcgtatcg gtctgggct ccgcgtctg ggctgcgtatg
9300

 gggtctgggg ctccgcgtc tggctgcga tgggtctgg ggctccgcgc tctggctgc
9360

 gatggggct gggctccgc gctctggct gcgatggggt ctggggctcc ggcgtctgg
9420

40 ctgcgtatgg gtctgggct ccgcgtctg ggctgcgtatg ggtctgggg ctccgcgtc
9480

 tggctgcga tgggtctgg ggctccgcgc tctggctgc gatggggct gggctctga
9540

45 gctctggct gcgatggggt ctggggcttg gtctagggcg cttctaattc cttgggttt
9600

 tcttggtctc tgcaggaagc tgcgttctcg aacactaggg tgagtgagcc tttaggaggg
9660

 cactggacaa gcccagagtc ctgggtcccc ggggttcgag ggtacatctg ctctggccct
9720

55 tcccatccca cacagccagg gagacccccc cagggtcagg cacgaggttg gcacacctaga
9780

gtctgccac ccaaaattcc tgggacattc gggaaatcct ttgtttacc attcctgcac
 9840

5 ctgccagccc gagtgagggt ctcctcgcc cttccacag cgaggcctcc ctccggctcc
 9900

ctcaggtgtc agctcgccc atcgccccct gcacctgtcc ccactcggtc tactcccccc
 9960

10 cacccactca ccaaggatgc tcagaggcct gcccagtcat tggaggagct gaggctgctc
 10020

tggagccccc aggctgccc gcagtggccg atgtggaagc tgcagagcct ggggagggag
 10080

15 ctctggccct ggccctctgcc ctaccctgca gcctccctga cctctgtcc tcttcccaagc
 10140

20 acagcagccc ccgggtctgc acctccagag cccaccctac caccagacac agagcccgaa
 10200

ccacctggac ctaccctcca gccatgaccc ttccctgtc ccaccacact gactccaaat
 10260

25 aaagagcttc tcccccaagct ctgggcagggc ctatctgtgg ggacggaggg gctcgacccg
 10320

30 gctccctggg aggccctgtc ggagggggag ccacaaggga agctggggag atgttggacc
 10380

tagcaagggc caaaccacaa gaggcacat gtccaaacagg ctgtgggct gcgtgatgtc
 10440

35 ctggaggggc ctccggaaga gccgcctcg gatctcaggc tcagcttggg acgggcaggg
 10500

ctctggcag gagcaattgg gaagccactg ggagggccgg agtggggaca cggcgtcaga
 10560

40 cctgcatcag tgggtccacc acagggtcctcc cgccccgggg gtctcagtgg gatcctccga
 10620

45 gcgggtccac tttancatcc tgngctcagt ggcacatcactg gata
 10664

<210> 16
 <211> 13591
 <212> ADN
 50 <213> Homo sapiens

<220>
 <223> Gène hOBPIIb

55 <400> 16
 tctctgttc tatattcgcc ctccaaatct gtgctcaatg cagcaactgg agtgacccct 60
 taaatacgta agtcacagct tgcccttgc agagctctcc agggtcttc actcagagca 120

gaagctgaag tcctcgttgt ggtccttaat ccctacatgg cagttccacc cactccccag 180
 5 cctcatgtgt ggccgggtctc cctggatcat ttgctgtggc tgctctgctg tgtgttccc 240
 gaaactgc当地 gcattctccc acctccaggc tggcaactgga tgctcctgcc acctggactc 300
 ctcttc当地 tgacgagctc atggcttgct tccttc当地 cttaaattcg gtgttgaat 360
 10 gccaccttgg cggggctgttt cctcatcaat tcatttaaag gacaataaac cccttgc当地 420
 gacactc当地 tccaacttgc当地 ccactttgc当地 ttttccatag cactgatcac catttaaaat 480
 aacgtatgaa ccagtgaccc atcaggatcc agagacaata gattgaacag ggattgcttgc 540
 15 cacaatgaac aatcagctct agaggtttaa agagaaccac aaagaatggc tcaggattgc 600
 acgagggcag ttaaggera aataaacagg ggtggggatc ctttccacaa gtctgggttc 660
 20 agacctcatg ggagaaggta tggttctccc aagggaaagct ggagaagttt gctgggttgc 720
 cccagccaca gctggccac ggtcaggca gaggccagca ggcaggaaa tgttaggctgg 780
 25 gtctggcaag caggaagcct ctctcccccac ccaaggcagg tgggctgggg ctgggaagcc 840
 tgcaagagtc actgagaacc cacaggtgca gatgccactg aatctcaata ggaagccatc 900
 tggggccggt cccctgaatt caatgtggta tggccgcca gacgtccccg acttgtgcca 960
 30 ctgccatting cataggera gaaagaaaaa ggaagaaatg gaagcatctg gaaccagtca
 1020
 tcctggaccc atgcttaggag gcgcttctcc ctacgc当地 accaacaag cttagcattg
 1080
 35 caccaggtgc acaagagaag cgccctgcttgc ctccagctcc attgc当地 cggccat
 1140
 gaagggtacg tggagctg gaggcaagac attgatagct ggc当地 gcaa ttc当地 tatt
 40 1200
 tattgttgc当地 attttaagtc ccctgc当地 agaatataag cccccc当地 gagc
 1260
 45 tggggatc atcgatgtat tccttgcc ccaaagaatg agaggcatct agaaagtc当地
 1320
 caaaaatcaa acataaaaaat gaacctt当地 tcagtc当地 ttatggat gggaaatttgc
 1380
 50 ggc当地 tccca agatggtaa gtaacaattg aacccttctg ctggatc当地 tggggatc
 1440
 tatctgttgc当地 aatgatgctt ctgaggag tggctgagtc tttggctcta ggtttttt
 55 1500
 ttttaagcat ttatgctt当地 tggggatc aaattaaaca ttaccctg agacattgta
 1560

gattcccatg cagttgtaaa aagccacaca gagctatcat gtgtatctc acctggctg
 1620

5 ctccagcccc aaccccagca atgacagacc ttttctccac tcctgaaatc tgctcattgc
 1680

10 aagaatgtcg tctattgcaa tcataaaatt gtgggattgg ctttttttt tcctgtgcag
 1740

15 catcattctc tggagattca tccattgtt gcatttatca atagtttatt ccattttact
 1800

20 tctgagtagt gctctatggt atggatgtac cacagtctgt ttaactattc acctgttgga
 1860

25 ggatgtctgt gtttatacat ttgggctatg acacatgtac aggttttgc atggacatca
 1920

30 gttttcattt ttctggaca aaggcccagg gtttctattt cggggttcta tgcttggc
 1980

35 agggttttt ttttaaacc tgtcaagcca tttccagaa actgtccagt ttccatgcat
 2040

40 cctcaccagg cttcagttatg atcactatga tcttatctca gccaccctaa taggtatgt
 2100

45 ctgatatac atggctttt tttgcatttc actgatgact aatgggtttt agcatcttt
 2160

50 catgtttta tttgccatct gtatatcctc tctagtcaag tgtcccttca tgtctttgt
 2220

55 ttacattcta ttttgaaac ttttgagttt taaaaattt tttataaaatc atagaaacta
 2280

attctttgtt gaatatgttag tttgtcaata ttttcttca gtctttggct tgtctttta
 2340

ttctgttaac agggcttctt acagagcaaa aggtttttat tttgatgaag tctatTTAA
 2400

caatTTTCC ttttatggat catatTTTGT gtgacacatc taaaatctcc tgacccagct
 2460

ccatatccca aagatTTCT ttctgttttca taaagctttt atagttttaa gttttatgtt
 2520

taagtctatg atccatTTTG agttaatTTT CTTGCAAGGGT gtgagaccaa ggTTGTGGTT
 2580

cttctttttt tttgggtttt ttgtctgtgg atgtccagtt aatccagccc atttgctaca
 2640

aaagctatct tccattgaat tgctttgca tctttgtaaa aacttaactg ggtatatgcg
 2700

tgcaggtctg tttgaggatt ctttattgtc tgtcccattg atctatgcat ctgtccatct
2760

5 gctagctata taatgagtct tgaaaggta gcctgttagct gggtatggtg gtgtgcacatct
2820

gtaatcccaag ctacttggga ggcggaggca ggagaatcac ttgaatctgg gaggtggagg
2880

10 ttgcaatgag tcgagattgt gctactgcac tccagcctgg gtgacagagc aagactctgt
2940

ctctaaaaaa aagaaagaag aagaaaagga aagaaagggt agcctgattc ctcccacttt
3000

15 attcttaaaa aaaaaagttt tagttattct agttcctttg cctttccata taaatgttag
3060

aataatcttgc tctatgtcta caaaaattcc ttctggaatt ttgatagaaa ttgtgttaaa
20 3120

tctttatatt atttgagaga aatgacgctt ttattatgtt gagtctccta atccatggat
3180

25 atagcacatc tctccatttgc tttagatctt ctttgattta ttttataatc attgcattat
3240

tttcagcata caaatccagc atatgggttg ttaaacttat gcctaggtat ctctttttt
3300

30 tagccactat aagtagtatt gtgttttaa gtttagggtc cgttactagt atgttagacac
3360

acaattgatc tttgtatatt tatcttgtat cctgcaacct tgctgaactc gcttaatagt
35 3420

tctaggaggt gtattgtttt ctttgtttt gttttcagt ttcttggat tttctacaga
3480

40 gacaatcatg tcatctgcag atgcagacag ttttcttcct tccttccaa tttgtatgtc
3540

ttaatttcc tttttaaaaaa acctttatttgc ctctgacttag aactttctgt actatgttaa
3600

45 atacaagtgg tgaaagtggc catccttgcc ttgtccctga tgtaaggag aaagcgtttgc
3660

taactgagta tcccagcctc aaaatgtgtc taaaaacttt tttccttct tgctttcagc
50 3720

cttgaacat acttcgaaac tctttatttc tcccttccc accaggcact tctgtgagca
3780

55 gtgctcgctt atctaattat gtgcttactt agaaattcca ggggccaatt ttgaaacaaa
3840

ccaggcagag agacccagct gcagaatcct ccctcttagg gggagttaca ggttagcctac
3900

5 cacttcccgg ctgaaatcag gatgactcaa accagaccc tggacagacg attaatgact
3960

catgataacc attggaacaa gatgcagacc aacatcctcc tgcaccattc ccacatattt
4020

10 cccacacctt ttccctccta aacccttcg ctcagtccag aaagtttcaa tggctttta
4080

aaggcatggg gcctggccat tcctctactg ctagcattt aataatgctg ctttccttc
4140

15 accacacttc acttctcatg ctttgacttc tgagcagcga gcagctggac ttgagccagt
4200

20 tacacatcag tcttcacca tgagaaatga tgtagccat aggtttttt tagatgctct
4260

ttgtcaagtt aaggaggttc tcttctattc ctactttctt gagagttgtt ttcctgaatg
4320

25 cgtgatgaat ttagtcaagt acctttctg cattgatcga tatgatcatg taatttttct
4380

tcttaacata ttaatatggt tgattttga gtattgtacc aggtttccat ccctggaata
4440

30 aactgctttt ggtcatggtg tagaattaat ttatatatatt attatttgc aatattttgt
4500

35 aaaggatttc tgtatctata ttcatgaggt atattttctt gcagcttcc tttttgtgtt
4560

gtcaggtttt gggtggaga tattctctca tcttctgaga ttatgcagag ttgggtgttat
4620

40 ttatttttta aatgggttgtt agaattttctt aatgaaacca tatggacatg aagaattatt
4680

tttggaaagct tttaaaaatta aaatgttaac tttaattgtc atagagctat taaagttatc
4740

45 tatcttatat tggcattgtt tttttgaga aattgggtca tttcacctat gttgtcaaatt
4800

50 atatgtggga aggtcttttc gtagtattcc cttattatcc tttgatgtct gcagggtctg
4860

tagtgcatagt ctctgttttgc ttccagatgt tggttatttg tgccttttttgc
4920

55 ttgtcagtc ttctagaac tgcctcaatt ttatataattt ttccagagaa ctaacccttg
4980

tttcattgtt tttctctgtt gttttctgc tttcgatTTT ttctgtttc tgctttgat
5040

5 ttattaattg ctgctcttat ctttgttatt tctttcttc ttgcttggg ttatTTTgc
5100

tcttctttt ctaggttctt gaagtggag cttagatagt tgatTTTgga cttctttt
5160

10 ctaatatatg catTAATgc ctctcagcac tgatTTAGct gtgtctcaca aatTTTgata
5220

tgttttgTT tgTTTgtTT tgaggtagag tctcgctctg tatcccaggc tggagtgcag
5280

15 tggcaCTATC ttggctcaCT gcaacCTCCA CCTCCTGGGT tgaagtaATT ctccTgCTTC
5340

20 agccccctaa gtagctggga ttacaggtat acgccaccat gcccagcaat ttatTTATT
5400

atttattttt gtatTTTtag tagagatgag gtttgtcat gctgtccagg ctggTctcga
5460

25 actcctgacc tcaagtgatc tgcccacgtg gggctccaaa gcactgagat tataggcata
5520

agccactgtg cccagcctgt tttgtTTTT tttgagatgg agtcttgctc tgcacCCAG
5580

30 gctggagtgc agtggcatga tagcccactg caacCTCCGC CTCCCGGTT caagtgattc
5640

tcctgcctca gcctcccgtat tagctggat tacaggcgtg caccaccaca ctcagctagt
5700

tttgcattt ttggtagaga tggTTTcg ccatgttggc caggctggc tcgaattcct
5760

40 gatcgcaagt gatccacCCG cctgggcctc ccaaagtgc ggaattatag gcacaaggca
5820

tcacGCCAG ctgatATGTT gtatTTcat tttcatCCAG CCTAGTATAT tttAAATT
5880

45 accttcaggc ttccTCTTTG acctgtggat tattcacagg tggTTTgttT ccaagcatct
5940

gaagaaAGAT tttccattat ctTCTGGCA ttgatttata gttAGATTCC atgtggTcag
6000

50 agactaccct ctatATCATT taaATTTA aaaATTCTT gaggctcatt ttctcatcca
6060

55 ggatATGGTC tatTTGGTA tatATTCTGG TGCACTTGAa aagaatgtgt gtatgctgt
6120

tgccaggtgg tgggttctac aaatgtcaat tcgattatgt taattgatgg gtggccgagt
6180

5 tctttgatgt ccctgctgg tttctgtcta gttgctctga gagaaggta ttgaagtctc
6240

caactataat tgggtttt taaatttctc ctttcagttc tgtttgctt cacatattgt
6300

10 gcagctaattt tttgggtcac acacgtttag gattactatg tcttcttgc agaatggccc
6360

acttattatt gtataatttc cctctctgtt aattagtctt tgttctaaag tctattttac
15 6420

ctgttattga aataatgctt ctgctttctt ttgattaatg tttatataatc tttttcatg
6480

20 ttttacttt caacctgtct ctattgttac atttgaagta agtttcttgt agacggcata
6540

tagttgtata atggctttta attcactttg ccaatcactg tcatttaata tttaaaccat
6600

25 ttacatttaa tataattatt aacatgttag agtttgtatt attttatag gggtgctata
6660

acaaaataact accaactggg tggcttagaa caaaaattta tttctcaca ttctggcgc
30 6720

cagaagtctg agatcaagtg agatttgaa ggtaaaaaa ccaacaaaat atgttaatga
6780

35 tgttaccatt gtggcaact gggactcagt cttctgagg gccatctgag aatagggtg
6840

gaaccatatt tgattgttcc acttagaagg aagagtccag accatttatac ccatcagttt
6900

40 ctatccctac agactgagga tgctgattca cttgcacatc tgggttcat cagtgtcagg
6960

ggaagcacag aagaaaaggt gcaggcactt aaggtggaa gctgtgaata cgtccgtgca
7020

45 ccaggaacga tctacaactg tggcaggtgg gccgacagcg tgggtcacca gtgccttcag
7080

50 ccacccttg cactgcactg acccactcac acagcaaattc aagtccattt tgtaactgat
7140

gcttcagcat ggtggccagc caccgtctt gaaataatta aaatagggtgg ttgggtggcac
7200

55 aggcgatggt ccctactgtg cagttggtct tgcagcctta actgagattc atcgtaagct
7260

ccccactctgc tccttcttgg cgtgcctcag tcccagccag cacatgctct ggtcttagtg
7320

5 cttgttttgtt ggtgccatct aaaccttctt atgcccaggt gtcgagtaaa gcccagacac
7380

10 tgtccacttc agtgaagcac ctaggaatta gcagccctga aaagctgtga tcactgtgag
7440

15 attcaaagaa aagggtcgta agggaaagtgg tgagagagaa agttgtgtgg tagcggcagc
7500

20 ccaacaggag ccaccatcag aaggaacaca gcatccaga gaagtcccc aggccacaaa
7560

25 gcactggggg ccagagatct caccctcat gtctgccatg gtttgcaccc caacccatcag
7620

30 ccctaccatc tcctgacggt gtctgacca gttgaaccca accagtatct aaaaggacac
7680

35 agttggcgaa tgagactggg acacagggca agatgggtgg atgatgggag actcctggag
7740

40 agcacccgtc gcactgagcg tggacctcat gggagagccc tttctctga agcatctgtg
7800

45 catgtcggtc cagcatttc ttcaaggatt gagccagcag caccagtgtt gtaaggtgt
7860

50 tagatcaatg attcacaaac aaccaatgaa atacgaggtg cttgggtac agtcaacaaa
7920

55 cacagagaac agagaaacct tcccgttgc gtagagactc cagcctgttt gctcatctct
7980

60 ggtcctgaag agcccaggcc ccatttttc ctgcccgggg ggatgagggt ggggggtttc
8040

65 tgcttggcag cacttgcgtt gggaggggtg agggactggg tcacgcccagc ccgttcgtta
8100

70 cagccccagg gtttattaga gtgtgcatta gtattattga gtaagtactt gagagtgtct
8160

75 gatcttggc cagagaggca caaagatgcc attgtgtggg tccagcacct gcctgcagag
8220

80 cgtgggtcag cttgggtcc cagggaaagat gacacacacc tggaaatgc agggtccggg
8280

85 agggagaagg caaagttctg gtcaggttg gctggggatg agccagcag agccaggtgc
8340

90 55 ccaagggacc tcagccacga actctgagca caggctggca ggtgactctt ggtgcctcat
8400

gcgacttgtt tatctaaagg gatgagattc aaggcctgtc ctgggggctg ggggcccgtca
8460

5 aagctgacga gagagggaa tgtatattct aacaggaga tggtcgatga atacgtgcag
8520

gaagaaaatgg acaggacaat ccagagagat cgcgtttct gaggacagca cagccaggct
8580

10 ccgttacgga gtgaagcggg gtcggggagg cggcggggtc cctcatatgg cccgaggagg
8640

ccgtatataa actgaccctg agccacaaaa tagtgcctc ctctgcctt aggaagctcg
8700

15 agtgcagggt ccaggtgggg aaaatcaatg cagagtgggt cccagagtgg gcgaaagctt
8760

20 gggctctagg gcgtgcggga ctcagctggc agcagccccca cgtctctctg gttctaaagc
8820

ccagtcatt tctgctcagc agggagatgg ccagttcccc agaggacttg cccagggtcc
8880

25 cagccgtggc cctgggaggg tgcaggggtt ggtgagggag tatcccgagg ctgtgggccc
8940

tgcctctgtg gcttgggtg cagatggaca ccaccctgcc ctgtcctctc gagtcctctg
9000

30 ggctcatctg agtggcgcg ttgggggtt gcaaactctg gatggcatcg tgtgtcttt
9060

35 cttctctggc tggcactgac ttgctgttca ccagatgcca gcaatgagga cactccccct
9120

cccgatgca ggactgggtt ccaccatggg gaggggtgca ccgtgccacg tgctcccagg
9180

40 acactggcca ggggctata aagaacatct cgagaggagc cagcacagcc ttgttcagac
9240

gcccagtgac ctgcccaggt cggcagcaca gagctctgga gatgaagacc ctgttcctgg
9300

45 gtgtcacgt cggcctggcc gctgcctgt cttcacccct ggaggaggag gatgtgagct
9360

50 gggttggcgt gggcgatgg aggagccagg cggactcctg ggcagggggc agtgcaggg
9420

gcccgtttt aggaggtgtc actgaaggct ggcttctgac cccgtgtcc cagcctgggg
9480

55 tggtggtgga ggggtcctat ttttcctcca gcagacacgg ccccccggagg cccaggccg
9540

gagcaggctc gtctcagggg ttccctgctgc actgacgcct gaagcccgaa ggtctcgca
9600

5 ggttggggccc tgtggaggga gggctcacct ggtgctgggg cccgggggtc catgggggtc
9660

agacatgccc tccttccact gggggctggg agccctgagc agggggctgg ctctaactca
9720

10 ctccagctga gctctaacta aggtgcagga acacagcctg ccttaaggg cagcagccgg
9780

gcaccatggg tgtctggta tagctgcagg cctgagtgcc agggtcagag tagaatctgg
9840

15 gccacccatg gtgggctcac ggccttggcc tgctccagat cacagggacc tggtacgtga
9900

20 aggccatggt ggtcgataag gactttccgg aggacaggag gcccaggaag gtgtccccag
9960

tgaaggtgac agccctggc ggtgggaagt tggaaagccac gttcaccttc atgtgagtgt
10020

25 tgcccactgc agggcccctc aggccacttt cgctcccctc cccagaccca cctggtgccc
10080

attgccccat ccacgtttcg ggtgttggga agagtcaccc cctgccttgg agggaaacag
10140

30 cctgggcata ctgaagctcg gtgggggtggg ggacagtggaa attttcaggt tgccgggtca
10200

gggccatgca ccaggtgagc tgaggatggg ccatgggtgt cctggagcc gtcggccgc
10260

tgtctcctgt tttccaggag ggaggatcggt tgcatccaga agaaaatcct gatgcggaaag
10320

40 acggaggagc ctggcaaata cagcgctgt gagccctcc cccactccca cccccacccct
10380

cccccacccgc caaccccaagt gcaccagcct ccacaggttag agagtgcacca ggctgcctt
10440

45 ttgccaggc cccagctctg cccacccctca aggaggggtc ggcctctcct tcctgggggg
10500

ctgggtggcc tgacatcaga caccgggtgt gacaggcttg tccgcagtcg agatggacca
10560

50 gatcagcgct gccctctggg aggccttagc cattgacaca ttgaggaagc tgaggattgg
10620

55 gacaaggagg ccaaggatta ggtacgggaa ggctgaggggt tagggatggg gaagctgagg
10680

gttaaggatg gggaaagctga gggtaagga tggggaggct gagggtagg gacggggcc
10740

5 aagggttagg gatcgggaag ctgagggtta gggatgagga ggccgagggt tagggatggg
10800

gaggccgagg gttagtgcgt cgaaagctga ggattagaga tggggaggct gaggggaaag
10860

10 atgagggtta gggacaggga agtggggcag gacagctgc tggagctggg aaggccggac
10920

tctagtccctg gacgtgctct gccttgtgg ctccattact tgcattggg cttcctgag
10980

15 aggaggctcc tgcctccgtg tccgggtcca cactgtgcgg agcagccagg cctggctcag
11040

20 gctgtccagg gcacctgggt gaccgctgaa acattcccgaa gtgtttcttc atgtggcccc
11100

gagtgttctc tctggaaatg acccattttc tctgtcatct acagatgggg gcaggaagct
11160

25 catgtacctg caggagctgc ccaggaggga ccactacatc ttttactgca aagaccagca
11220

ccatggggc ctgctccaca tggaaagct tgtgggtgag gggcttgctg gggctgcgt
11280

30 gtcctgcccc acggctcttg cctctggaaag ccgggtggaaac cagcaccacc atgccctggc
11340

atggggaaaa aggaagcccc ctgtgcggc tttcatgtgc cgggcaccgc tgggcagccg
11400

gtccaagtac ctgaagtggg aatggatgg gccaggcagg gacagacatg gccctcagtg
11460

40 acatgaacct gccaagggcc acttctggg tctcgggtgt aggagggtca ctttaagggg
11520

gagggtcacc ttaagcgtcc taagagagct cttcatggg caggaccct ccaggcagg
11580

45 agggtggccg gtgcctctgg gagacaaggt ccctccaggt gagggtgtg accctgcagg
11640

50 gtggcattgg gagctgccc ggtccctgg gttgccag tggcttggag agtccccgcc
11700

actgtcccccc ttctgggaa cctctcatct gggcagtgac tgtctcctct gtccccagtc
11760

55 ccacccctga gctctcatcc attctaagt ctccctccct gttgtccgg tgctggctca
11820

tggccatctc ttctgtcccc ggccccaccc cagagctctg gtccactctt gggctctcc
 11880

5 ctttgtcctg gtgctggca gtggccatct cctctgtccc cagccccacc cccgagctct
 11940

gatccactct caggcctctc cccctgtcct ggtgctggc cgtggcgtc tccttcggc
 12000

10 atctgctccc tgcagggccg tgccgctgtc cccacgtcgg ctcacctggc cacctcacct
 12060

gcaggttagga attctgatac caaccggag gccctggaag aatttaagaa attggtgca
 12120

15 cgcaaggagac tctcgagga ggacatttc acgccccctgc agacgggtga ggatggctgt
 12180

20 gcccagtccc ctgtgtccct ctgctgtgtc tgtctgctat ctccagtgtc ccatgacccc
 12240

catgtccctcc catgtcccccc gcattccca tgtgccccga gtctcctcgc aggggctccc
 12300

25 gggccctgtt tagcgccctc ctcattggag gctctgtgct ctgggctgca atggggctcg
 12360

gggctcccgcg ctctggctg cgatggggtc tggggctccg cactctggc tgcgatgggg
 12420

30 tctggggctc cgcgctctgg gctgcgatgg gctctggggc tctgagctct gggctgtgat
 12480

ggggtctggg ccctggctta gggcgcccttc taatccctgg gtttttcttg gtctctgcag
 12540

gaagctgcgt tcccgAACAC tagggtgagt gagccttag gagggcactg gacaagccca
 12600

40 gagtcctggg ttcccggtt tcgagggtac atctgctctg gcccctccca tcccacacag
 12660

ccagggagac ccccccaggg tcaggcacga gttggcacc tcagagtctg cccacccaca
 12720

45 gttcctggga cattcgaa gtccttgtt ttatcattcc tgcacctgcc agcccgagtg
 12780

agggtccctcc tcggccttcc cacagcgagg cctccctccg gtcacccatcg gtgtcagctc
 12840

agcccatcgc cccctgcacc tgccccacg cggccactc cccaccatcc acccaccacaa
 12900

55 gatgctcaga ggcctgcccc gtcattggag gagctgaggt cgctctggag ccccgaggtt
 12960

gcccagcagt ggccgatgtg ggggctgcag agcctgggga gggagctctg gccctggcct
13020

5 ctgccctacc ctgcagcctc cctgacacct ctgcctcttc ccagcacagc agccccccggg
13080

tctgcacctc cagagccac cctaccacca gacacagagc ccggaccacc tggacctacc
13140

10 ctccagccat gacccttccc tgctcccacc cacctgactc caaataaaagt ccttctcccc
13200

cagctctggg caggcctatc tgtggggaca gaggggcttg caccgcctcc ctgggaggtc
13260

15 tgctgggagg gggagccaca aggaaagctg gggagatgtt ggacctagca agggccaaac
13320

20 cacaagaggg acgatgtcca acaggcctga ggggctgcgt gatgtcctgg aggggcctcc
13380

ggaagagccg ccctcaggc tcaggttcag cttaggacag gcagggccct gggcaggagc
13440

25 gtttggaaag ccactgggga ggacaggagc ggggacacgg cgtcaggcgt gcagcagtgg
13500

ggccaccgca gggctccgc ctcagggc tcagagggat cctccaagct ccccattt
13560

30 agcagcctgg gctcagtggc agcactggag a
13591

REVENDICATIONS

1. Polypeptide isolé comprenant une séquence en amino-acides ayant au moins 90% d'identité avec les séquences en amino-acides SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14.
2. Polypeptide isolé caractérisé en ce qu'il comprend un polypeptide choisi parmi :
 - 10 a) un polypeptide de séquence SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14;
 - b) un polypeptide variant de polypeptide de séquences d'acides aminés défini en a) ;
 - c) un polypeptide homologue au polypeptide défini en a) ou b) et comportant au moins 90 % d'identité avec ledit polypeptide de a) ;
 - 15 d) un fragment d'au moins 15 acides aminés consécutifs d'un polypeptide défini en a), b) ou c) ;
 - e) un fragment biologiquement actif d'un polypeptide défini en a), b) ou c).
- 20 3. Polypeptide isolé sélectionné parmi un polypeptide correspondant à la séquence SEQ ID N°2 et dénommé OBPII_{aα}, à la séquence SEQ ID N°4 et dénommée OBPII_{aβ}, à la séquence SEQ ID N°6 et dénommée OBPII_{aγ}, à la séquence SEQ ID N°10 et dénommée OBPII_{bα}, à la séquence SEQ ID N°12 et dénommée OBPII_{bβ}.
4. Polypeptide selon l'une quelconque des revendications 1 à 3 caractérisé en ce qu'il comporte au moins le domaine Gly-Thr-Trp-Tyr.
- 30 5. Polynucléotide isolé caractérisé en ce qu'il code pour un polypeptide selon l'une quelconque des revendications 1 à 4.

6. Polynucléotide selon la revendication 5 de séquence SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13.

5

7. Polynucléotide isolé caractérisé en ce qu'il comprend un polynucléotide choisi parmi :

- a) un polynucléotide de séquence SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13
10 ou dont la séquence est celle de l'ARN correspondant à la séquence SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13 ;
- b) un polynucléotide dont la séquence est complémentaire de la séquence d'un polynucléotide défini en a),
15
- c) un polynucléotide dont la séquence comporte au moins 80% d'identité, avec un polynucléotide défini en a) ou b),
- d) un polynucléotide hybride dans des conditions de forte stringence avec une séquence de polynucléotide défini en a), b) ou c),
- e) un fragment d'au moins 15 nucléotides consécutifs, de
20 préférence 21 nucléotides consécutifs, et de manière préférée 30 nucléotides consécutifs d'un polynucléotide défini en a), b), c) ou d).

8. Polynucléotide selon la revendication 7 caractérisé en ce qu'il est marqué directement ou indirectement par un composé radioactif ou
25 un composé non radioactif.

9. Utilisation d'un polynucléotide selon la revendication 8 en tant que sonde pour la détection de séquences nucléiques.

10. Utilisation d'un polynucléotide selon l'une quelconque des revendications 5 à 8 en tant qu'amorce pour l'amplification ou la polymérisation de séquences nucléiques.
- 5 11. Utilisation d'un polynucléotide selon la revendication 8 en tant que oligonucléotide sens ou antisens pour contrôler l'expression du produit protéique correspondant.
- 10 12. Vecteur recombinant de clonage d'un polynucléotide selon l'une des revendications 5 à 8 et/ou d'expression d'un polypeptide selon l'une des revendications 1 à 4 caractérisé en ce qu'il contient un polynucléotide selon l'une quelconque des revendications 5 à 8.
- 15 13. Vecteur selon la revendication 12 caractérisé en ce qu'il comporte les éléments permettant l'expression éventuellement la sécrétion dudit polypeptide dans une cellule hôte.
- 20 14. Vecteur selon l'une quelconque des revendications 12 à 13 caractérisé en ce que les éléments permettant l'expression dudit polypeptide sont choisis parmi :
- 25 a) le polynucléotide isolé de séquence SED ID N°15 et SEQ ID N°16;
- b) un polynucléotide dont la séquence est complémentaire de la séquence du polynucléotide défini en a) ;
- c) un polynucléotide dont la séquence comporte au moins 80% d'identité avec un polynucléotide défini en a) ou en b) ;
- d) un polynucléotide hybride dans des conditions de forte stringence avec une séquence du polynucléotide défini en a), b) ou c).
- 30 15. Vecteur selon les revendications 13 et 14 pour l'expression dans les cellules eucaryotes sélectionnés parmi l'ADN viral et l'ADN nu.

16. Cellule hôte, caractérisée en ce qu'elle est transformée par un vecteur selon l'une des revendications 12 à 15.

5 17. Procédé de préparation d'un polypeptide recombinant caractérisé en ce que l'on cultive une cellule hôte selon la revendication 16 dans des conditions permettant l'expression et éventuellement la sécrétion dudit polypeptide recombinant et que l'on récupère ledit polypeptide recombinant.

10

18. Polypeptide recombinant susceptible d'être obtenu par un procédé selon la revendication 17.

15 19. Utilisation d'un polypeptide choisi parmi un polypeptide selon l'une quelconque des revendications 1 à 4 et 18 ou l'un de ses fragments, LCN1, RBP et lapolipoprotéine D comme protéine de liaison à un ligand hydrophobe de préférence une molécule odorante.

20 20. Utilisation d'un polypeptide selon l'une quelconque des revendications 1 à 4 et 18 ou l'un de ses fragments, LCN1, RBP et lapolipoprotéine D comme inhibiteur compétitif comme, agoniste ou antagoniste des récepteurs cellulaires aux lipocalines.

25 21. Anticorps monoclonal ou polyclonal et ses fragments caractérisé en ce qu'il est dirigé spécifiquement contre un polypeptide isolé selon l'une des revendications 1 à 4 et 18.

30 22. Utilisation d'un anticorps selon la revendication 21 pour la mise en évidence de la présence d'un polypeptide selon l'une des revendications 1 à 4 et 18 dans un échantillon biologique.

23. Procédé de détection d'anticorps dirigé contre hOBPII dans le sérum humain de patient allergique et/ou asthmatique en utilisant un polypeptide hOBPII.

5 24. Procédé pour contrôler la volatilisation d'une odeur caractérisé en ce qu'il comprend une étape de liaison de ladite odeur avec un polypeptide selon l'une quelconque des revendications 1 à 4 et 18, ou LCN1, RBP et l'apolipoprotéine D.

10 25. Procédé selon la revendication 24 caractérisé en ce que le polypeptide est lié à un support solide.

26. Procédé selon la revendication 24 caractérisé en ce que le polypeptide est dans une composition liquide.

15 27. Procédé selon la revendication 26 caractérisé en ce que ladite composition est une composition parfumée pour la peau.

28. Procédé de criblage de molécule, de préférence les odeurs
20 ou saveurs qui comprend de passer la molécule sur un substrat qui comprend un polypeptide selon l'une quelconque des revendications 1 à 4 et 19, LCN1, RBP et l'apolipoprotéine D lié audit substrat, ledit polypeptide liant ladite odeur ou saveur et récupérer ladite odeur ou saveur depuis le polypeptide si nécessaire.

25 29. Procédé selon la revendication 28 caractérisé en ce que les molécules sont des phéromones humaines.

30 30. Procédé pour solubiliser des molécules lipophiles caractérisé en ce qu'il comprend la liaison de ladite molécule à un

polypeptide selon l'une quelconques des revendications 1 à 4 et 19, LCN1, RBP ou l'apolipoprotéine D.

31. Application des polypeptides selon l'une des revendications
5 1 à 4 et 19, LCN1, RBP ou l'apolipoprotéine D en combinaison avec des acides gras alimentaires, à titre d'additif alimentaire.

32. Application selon la revendication 31 dans le traitement des hyperlipidémies et de l'obésité.

10

33. Application selon les revendications 30 pour complémenter les laits non-maternels.

34. Polypeptide selon l'une quelconque des revendications 1 à 4
15 et 19 en tant qu'agent de ciblage de composé pharmaceutique.

35. Polypeptide selon la revendication 34 caractérisé en ce que ledit polypeptide est exprimé sous la forme d'une protéine de fusion avec une protéine permettant un adressage cellulaire spécifique.

20

36. Polypeptide selon la revendication 35 caractérisé en ce que ladite protéine permettant un adressage cellulaire spécifique est choisie dans le groupe composé des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps.

37. Polypeptide selon la revendication 34 caractérisé en ce que ledit polypeptide est associé à une molécule permettant un adressage cellulaire spécifique.

30

38. Polypeptide selon la revendication 37 caractérisé en ce que ladite molécule permettant un adressage cellulaire spécifique est choisie dans le groupe composé des stéroïdes, des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps.

39. Polypeptide selon l'une quelconque des revendications 1 à 4 et 19 en tant que transporteur de composé pharmaceutique.

10 40. Composition pharmaceutique comprenant un composé pharmaceutique lié au moins à un polypeptide selon l'une des revendications 34 à 39 et un véhicule pharmaceutiquement acceptable.

15 41. Composition pharmaceutique selon la revendication 40 caractérisée en ce que le composé pharmaceutique est choisi dans le groupe des agents anticancéreux.

20 42. Composition pharmaceutique selon la revendication 41 caractérisée en ce que ledit agent anticancéreux est un isotope radioactif choisi parmi le groupe : l'Iode¹³¹, l'Yttrium⁹⁰, l'Or¹⁹⁹, le Palladium¹⁰⁰, le Cuivre⁶⁷, le Bismuth²¹⁷ et l'Antimoine²¹¹.

25 43. Composition pharmaceutique selon l'une quelconque des revendications 40 à 42 caractérisée en ce que ledit polypeptide selon l'une quelconque des revendications 34 à 39 constitue une forme retard de délivrance dudit composé pharmaceutique dans l'organisme.

30 44. Composition pharmaceutique comprenant un vecteur d'expression selon les revendications 12 ou 13 et un véhicule pharmaceutiquement acceptable.

45. Composition pharmaceutique selon l'une quelconque des revendications 36 à 42 pour le traitement du cancer choisi de préférence parmi le cancer du sein, le cancer de l'utérus, le cancer de la prostate, le cancer du foie, du carcinome des cellules épithéliales pulmonaires.

5
ORIGINAL

CABINET REGIMBEAU
CONSEILS EN PROPRIETE INDUSTRIELLE
26, Avenue Kléber
75116 PARIS

REVENDICATIONS

1. Polypeptide isolé comprenant une séquence en amino-acides ayant au moins 90% d'identité avec les séquences en amino-acides SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ 5 ID N°12 ou SEQ ID N°14.

2. Polypeptide isolé caractérisé en ce qu'il comprend un polypeptide choisi parmi :

- 10 a) un polypeptide de séquence SEQ ID N°2, SEQ ID N°4, SEQ ID N°6, SEQ ID N°8, SEQ ID N°10, SEQ ID N°12 ou SEQ ID N°14;
- b) un polypeptide variant de polypeptide de séquences d'acides aminés défini en a) ;
- c) un polypeptide homologue au polypeptide défini en a) ou b) et 15 comportant au moins 90 % d'identité avec ludit polypeptide de a) ;
- d) un fragment d'au moins 15 acides aminés consécutifs d'un polypeptide défini en a), b) ou c) ;
- e) un fragment biologiquement actif d'un polypeptide défini en a), b) ou c).

20 3. Polypeptide isolé sélectionné parmi un polypeptide correspondant à la séquence SEQ ID N°2 et dénommé OBPII_{αα}, à la séquence SEQ ID N°4 et dénommé OBPII_{αβ}, à la séquence SEQ ID N°6 et dénommé OBPII_{αγ}, à la séquence SEQ ID N°10 et dénommé OBPII_{βα}, 25 à la séquence SEQ ID N°12 et dénommé OBPII_{ββ}.

4. Polypeptide selon l'une quelconque des revendications 1 à 3 caractérisé en ce qu'il comporte au moins le domaine Gly-Thr-Trp-Tyr.

30 5. Polynucléotide isolé caractérisé en ce qu'il code pour un polypeptide selon l'une quelconque des revendications 1 à 4.

6. Polynucléotide selon la revendication 5 de séquence SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13.

5

7. Polynucléotide isolé caractérisé en ce qu'il comprend un polynucléotide choisi parmi :

- a) un polynucléotide de séquence SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13
10 ou dont la séquence est celle de l'ARN correspondant à la séquence SEQ ID N° 1, SEQ ID N° 3 , SEQ ID N° 5, SEQ ID N°7, SEQ ID N°9, SEQ ID N°11 ou SEQ ID N° 13 ;
- b) un polynucléotide dont la séquence est complémentaire de la séquence d'un polynucléotide défini en a),
15 c) un polynucléotide dont la séquence comporte au moins 80% d'identité, avec un polynucléotide défini en a) ou b),
d) un polynucléotide hybride dans des conditions de forte stringence avec une séquence de polynucléotide défini en a), b) ou c),
e) un fragment d'au moins 15 nucléotides consécutifs, de
20 préférence 21 nucléotides consécutifs, et de manière préférée 30 nucléotides consécutifs d'un polynucléotide défini en a), b), c) ou d).

8. Polynucléotide selon la revendication 7 caractérisé en ce qu'il est marqué directement ou indirectement par un composé radioactif ou
25 un composé non radioactif.

9. Utilisation d'un polynucléotide selon la revendication 8 en tant que sonde pour la détection de séquences nucléiques.

10. Utilisation d'un polynucléotide selon l'une quelconque des revendications 5 à 8 en tant qu'amorce pour l'amplification ou la polymérisation de séquences nucléiques.

5 11. Utilisation d'un polynucléotide selon la revendication 8 en tant que oligonucléotide sens ou antisens pour contrôler l'expression du produit protéique correspondant.

10 12. Vecteur recombinant de clonage d'un polynucléotide selon l'une des revendications 5 à 8 et/ou d'expression d'un polypeptide selon l'une des revendications 1 à 4 caractérisé en ce qu'il contient un polynucléotide selon l'une quelconque des revendications 5 à 8.

15 13. Vecteur selon la revendication 12 caractérisé en ce qu'il comporte les éléments permettant l'expression éventuellement la sécrétion dudit polypeptide dans une cellule hôte.

20 14. Vecteur selon l'une quelconque des revendications 12 à 13 caractérisé en ce que les éléments permettant l'expression dudit polypeptide sont choisis parmi :

a) le polynucléotide isolé de séquence SED ID N°15 et SEQ ID N°16;

b) un polynucléotide dont la séquence est complémentaire de la séquence du polynucléotide défini en a) ;

25 c) un polynucléotide dont la séquence comporte au moins 80% d'identité avec un polynucléotide défini en a) ou en b) ;

d) un polynucléotide hybride dans des conditions de forte stringence avec une séquence du polynucléotide défini en a), b) ou c).

30 15. Vecteur selon les revendications 13 et 14 pour l'expression dans les cellules eucaryotes sélectionnés parmi l'ADN viral et l'ADN nu.

16. Cellule hôte, caractérisée en ce qu'elle est transformée par un vecteur selon l'une des revendications 12 à 15.
- 5 17. Procédé de préparation d'un polypeptide recombinant caractérisé en ce que l'on cultive une cellule hôte selon la revendication 16 dans des conditions permettant l'expression et éventuellement la sécrétion dudit polypeptide recombinant et que l'on récupère ledit polypeptide recombinant.
- 10 18. Polypeptide recombinant susceptible d'être obtenu par un procédé selon la revendication 17.
- 15 19. Utilisation d'un polypeptide choisi parmi un polypeptide selon l'une quelconque des revendications 1 à 4 et 18 ou l'un de ses fragments, comme protéine de liaison à un ligand hydrophobe de préférence une molécule odorante.
- 20 20. Utilisation d'un polypeptide selon l'une quelconque des revendications 1 à 4 et 18 ou l'un de ses fragments, comme inhibiteur compétitif comme, agoniste ou antagoniste des récepteurs cellulaires aux lipocalines.
- 25 21. Anticorps monoclonal ou polyclonal et ses fragments caractérisé en ce qu'il est dirigé spécifiquement contre un polypeptide isolé selon l'une des revendications 1 à 4 et 18.
- 30 22. Utilisation d'un anticorps selon la revendication 21 pour la mise en évidence de la présence d'un polypeptide selon l'une des revendications 1 à 4 et 18 dans un échantillon biologique.

fe
rectificative

23. Procédé de détection d'anticorps dirigé contre hOBPII dans le sérum humain de patient allergique et/ou asthmatique en utilisant un polypeptide hOBPII.

5 24. Procédé pour contrôler la volatilisation d'une odeur caractérisé en ce qu'il comprend une étape de liaison de ladite odeur avec un polypeptide selon l'une quelconque des revendications 1 à 4 et 18.

10 25. Procédé selon la revendication 24 caractérisé en ce que le polypeptide est lié à un support solide.

26. Procédé selon la revendication 24 caractérisé en ce que le polypeptide est dans une composition liquide.

15 27. Procédé selon la revendication 26 caractérisé en ce que ladite composition est une composition parfumée pour la peau.

20 28. Procédé de criblage de molécule, de préférence les odeurs ou saveurs qui comprend de passer la molécule sur un substrat qui comprend un polypeptide selon l'une quelconque des revendications 1 à 4 et 18 lié audit substrat, ledit polypeptide liant ladite odeur ou saveur et récupérer ladite odeur ou saveur depuis le polypeptide si nécessaire.

25 29. Procédé selon la revendication 28 caractérisé en ce que les molécules sont des phéromones humaines.

30 30. Procédé pour solubiliser des molécules lipophiles caractérisé en ce qu'il comprend la liaison de ladite molécule à un polypeptide selon l'une quelconques des revendications 1 à 4 et 18.

31. Application des polypeptides selon l'une des revendications 1 à 4 et 18 en combinaison avec des acides gras alimentaires, à titre d'additif alimentaire.

5

32. Application selon la revendication 31 pour la préparation d'un médicament destiné au traitement des hyperlipidémies et de l'obésité.

10

33. Application selon la revendication 31 pour complémenter les laits non-maternels.

34. Polypeptide selon l'une quelconque des revendications 1 à 4 et 18 en tant qu'agent de ciblage de composé pharmaceutique.

15

35. Polypeptide selon la revendication 34 caractérisé en ce que ledit polypeptide est exprimé sous la forme d'une protéine de fusion avec une protéine permettant un adressage cellulaire spécifique.

20

36. Polypeptide selon la revendication 35 caractérisé en ce que ladite protéine permettant un adressage cellulaire spécifique est choisie dans le groupe composé des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps.

25

37. Polypeptide selon la revendication 34 caractérisé en ce que ledit polypeptide est associé à une molécule permettant un adressage cellulaire spécifique.

30

38. Polypeptide selon la revendication 37 caractérisé en ce que ladite molécule permettant un adressage cellulaire spécifique est

fond de recherche

choisie dans le groupe composé des stéroïdes, des interleukines, des cytokines, des lymphokines, des interférons, des facteurs de croissance, des hormones, des anticorps.

5 39. Polypeptide selon l'une quelconque des revendications 1 à 4 et 18 en tant que transporteur de composé pharmaceutique.

10 40. Composition pharmaceutique comprenant un composé pharmaceutique lié au moins à un polypeptide selon l'une des revendications 34 à 39 et un véhicule pharmaceutiquement acceptable.

15 41. Composition pharmaceutique selon la revendication 40 caractérisée en ce que le composé pharmaceutique est choisi dans le groupe des agents anticancéreux.

20 42. Composition pharmaceutique selon la revendication 41 caractérisée en ce que ledit agent anticancéreux est un isotope radioactif choisi parmi le groupe : l'Iode¹³¹, l'Yttrium⁹⁰, l'Or¹⁹⁹, le Palladium¹⁰⁰, le Cuivre⁶⁷, le Bismuth²¹⁷ et l'Antimoine²¹¹.

25 43. Composition pharmaceutique selon l'une quelconque des revendications 40 à 42 caractérisée en ce que ledit polypeptide selon l'une quelconque des revendications 34 à 39 constitue une forme retard de délivrance dudit composé pharmaceutique dans l'organisme.

30 44. Composition pharmaceutique comprenant un vecteur d'expression selon les revendications 12 ou 13 et un véhicule pharmaceutiquement acceptable.

30 45. Composition pharmaceutique selon l'une quelconque des revendications 40 à 44 pour le traitement du cancer choisi de

préférence parmi le cancer du sein, le cancer de l'utérus, le cancer de la prostate, le cancer du foie, du carcinome des cellules épithéliales pulmonaires.

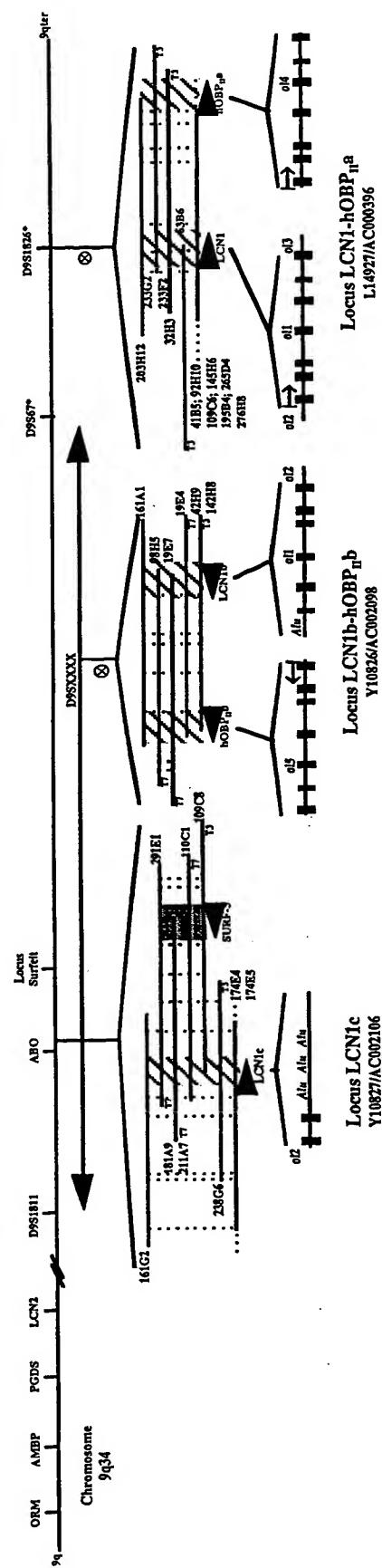


FIG. 1

DET REGIMBEAU

ORIGINAL

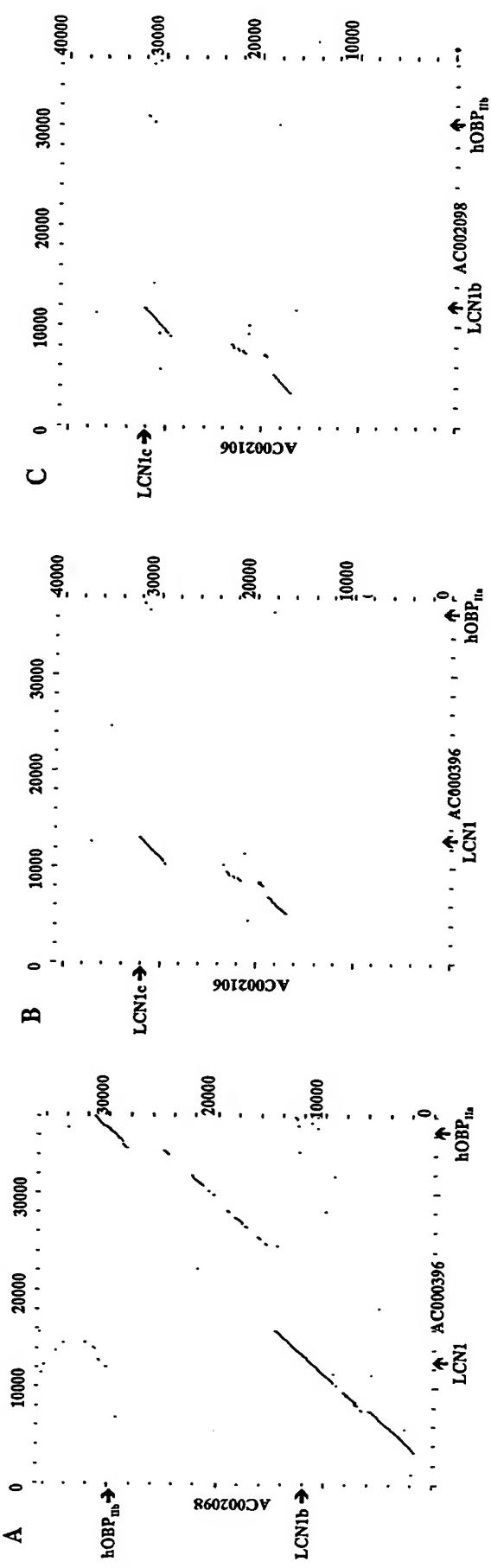


FIG. 2

FIG. 3

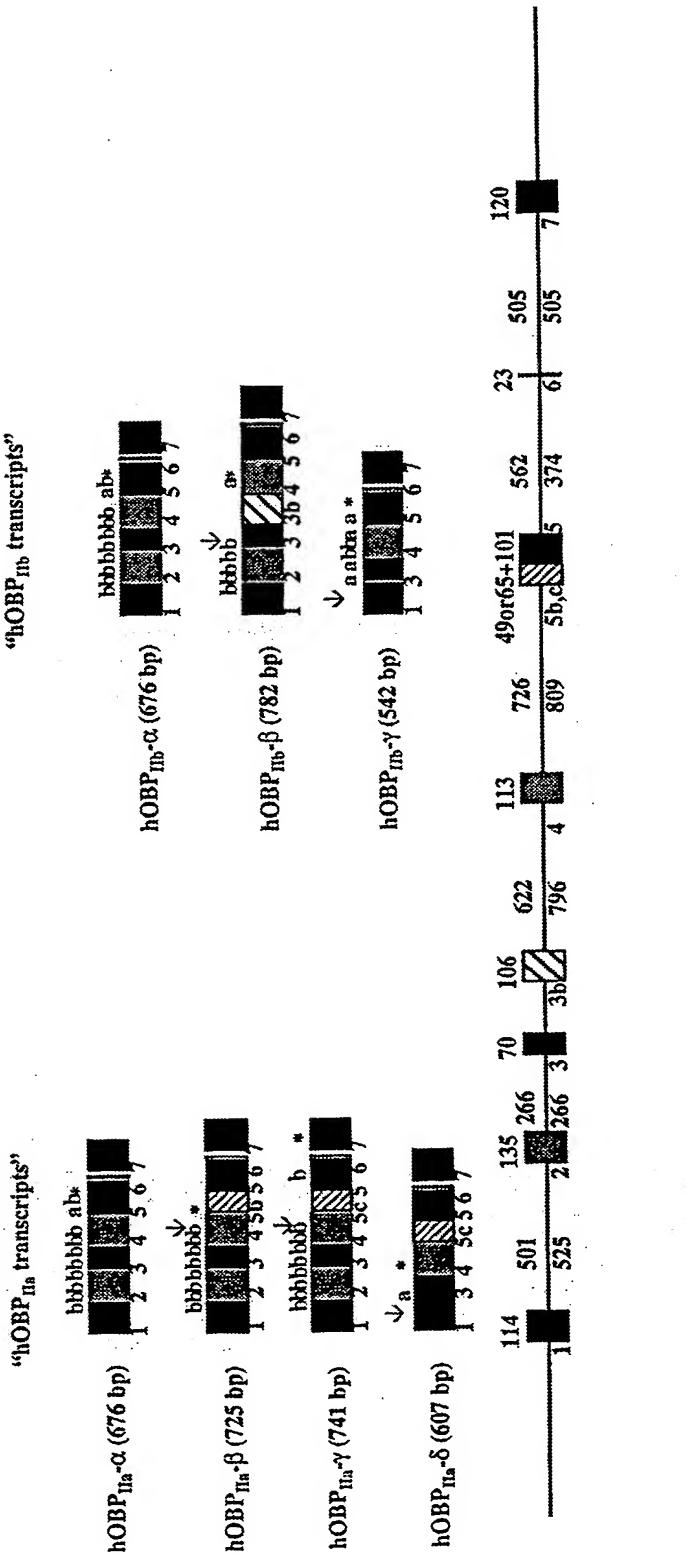
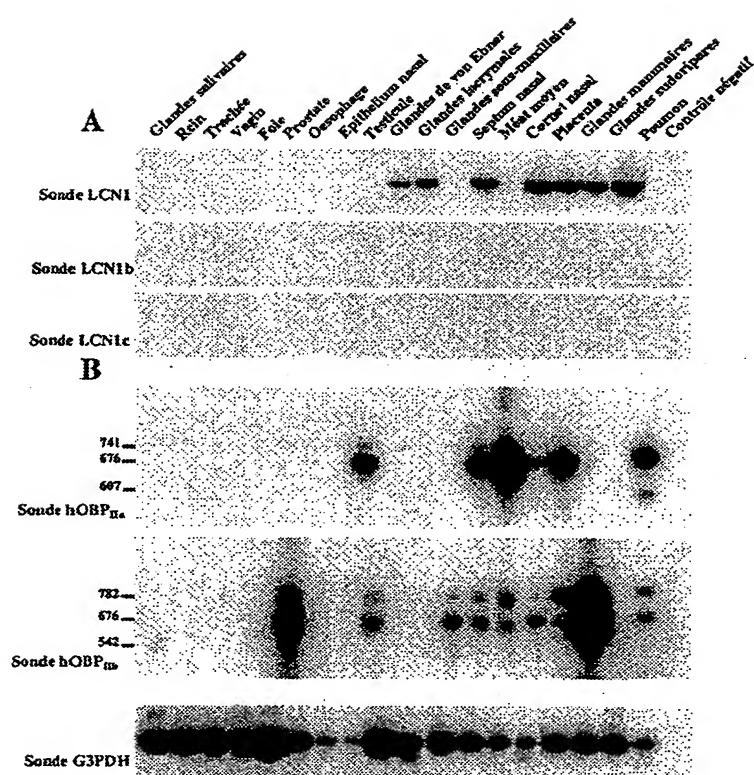


FIG. 4

FIG.5

FIG.6

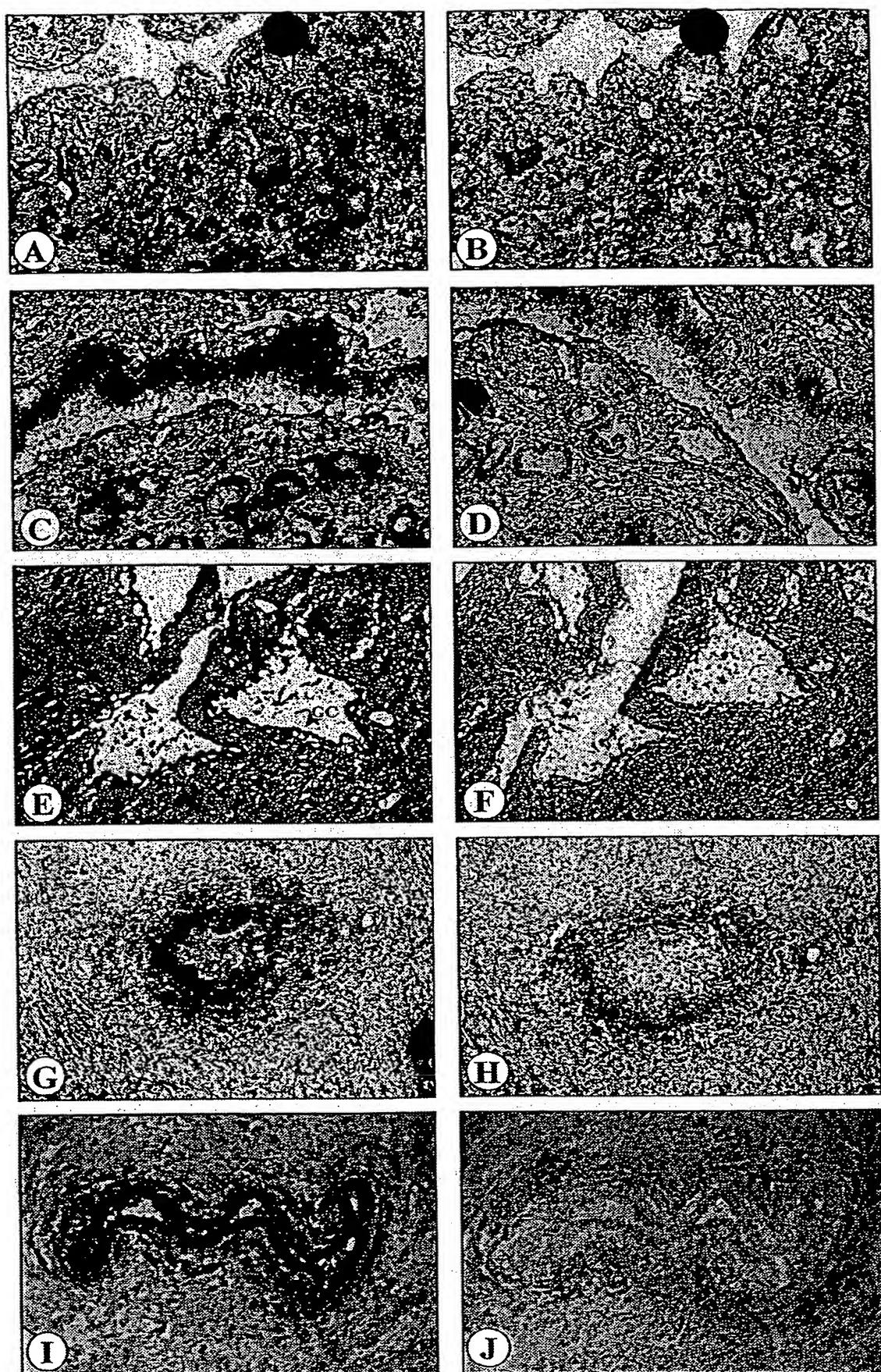


FIG. 7

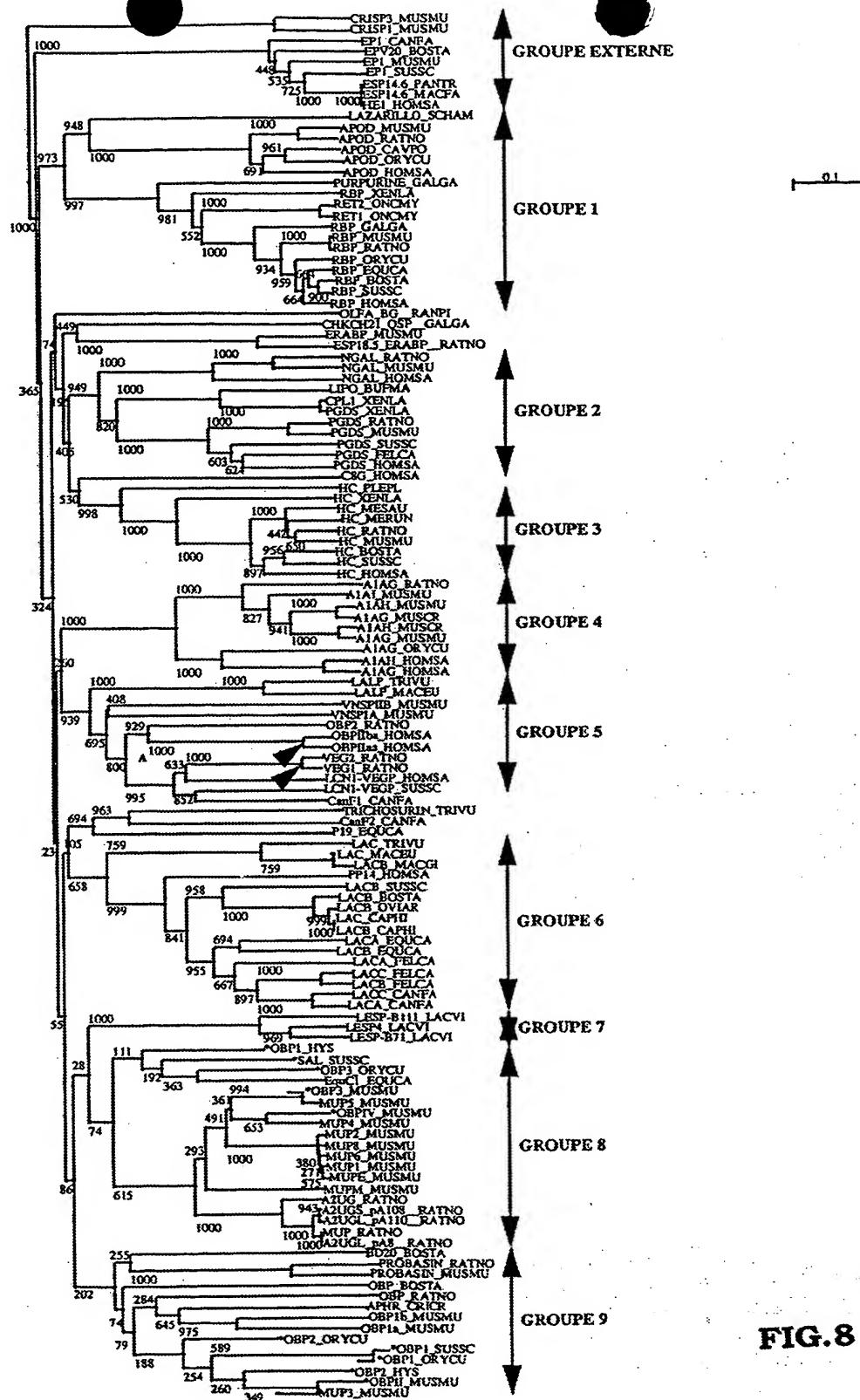


FIG. 8

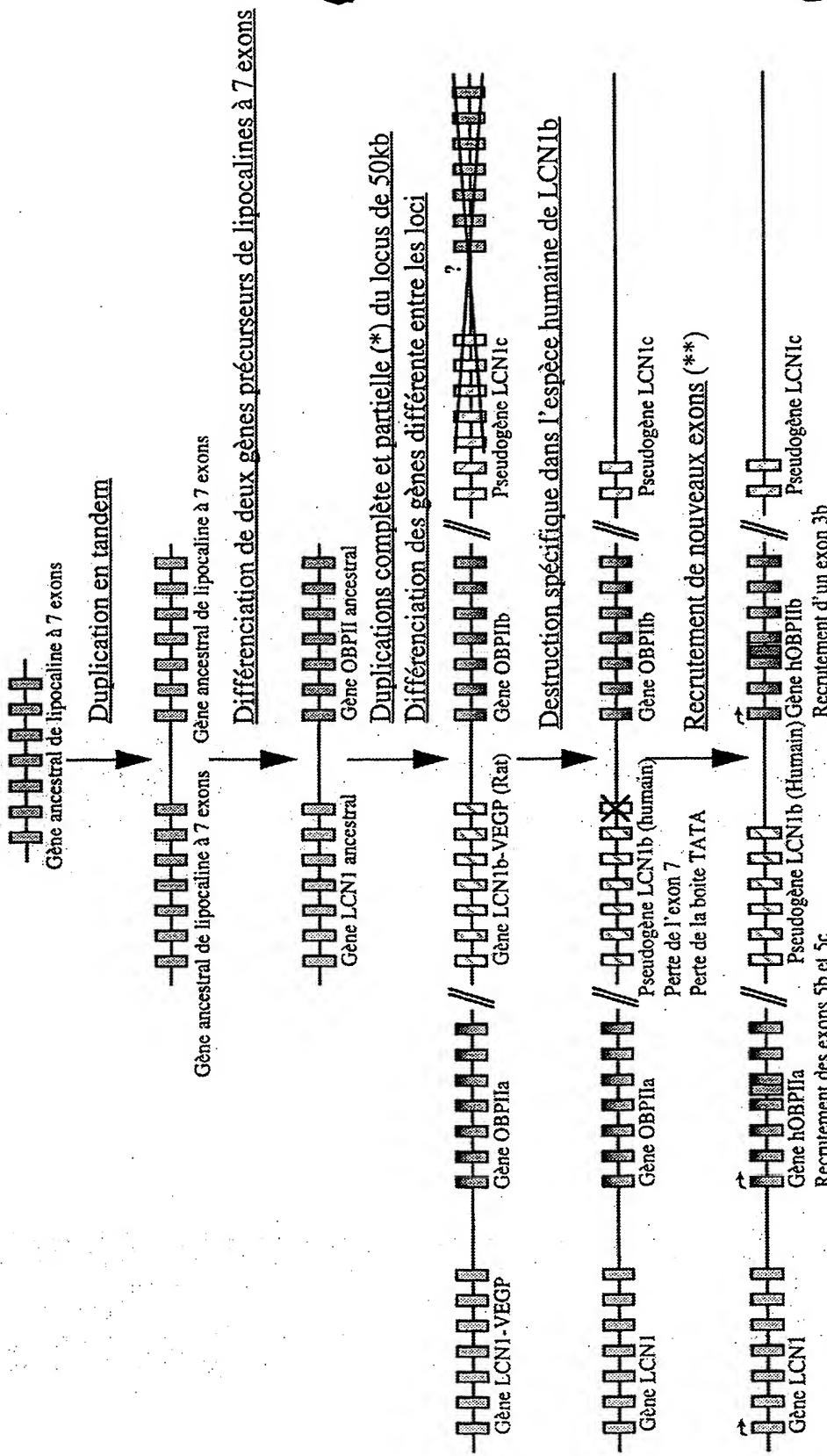
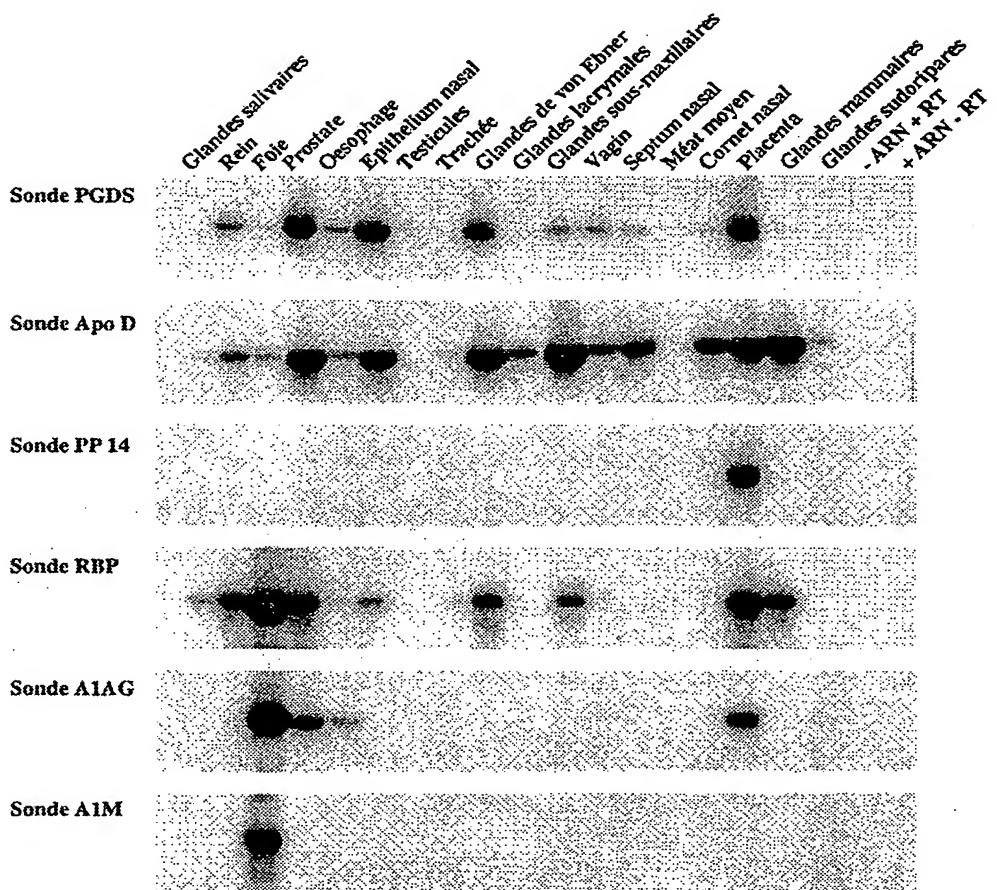


FIG. 9

**FIG. 10**

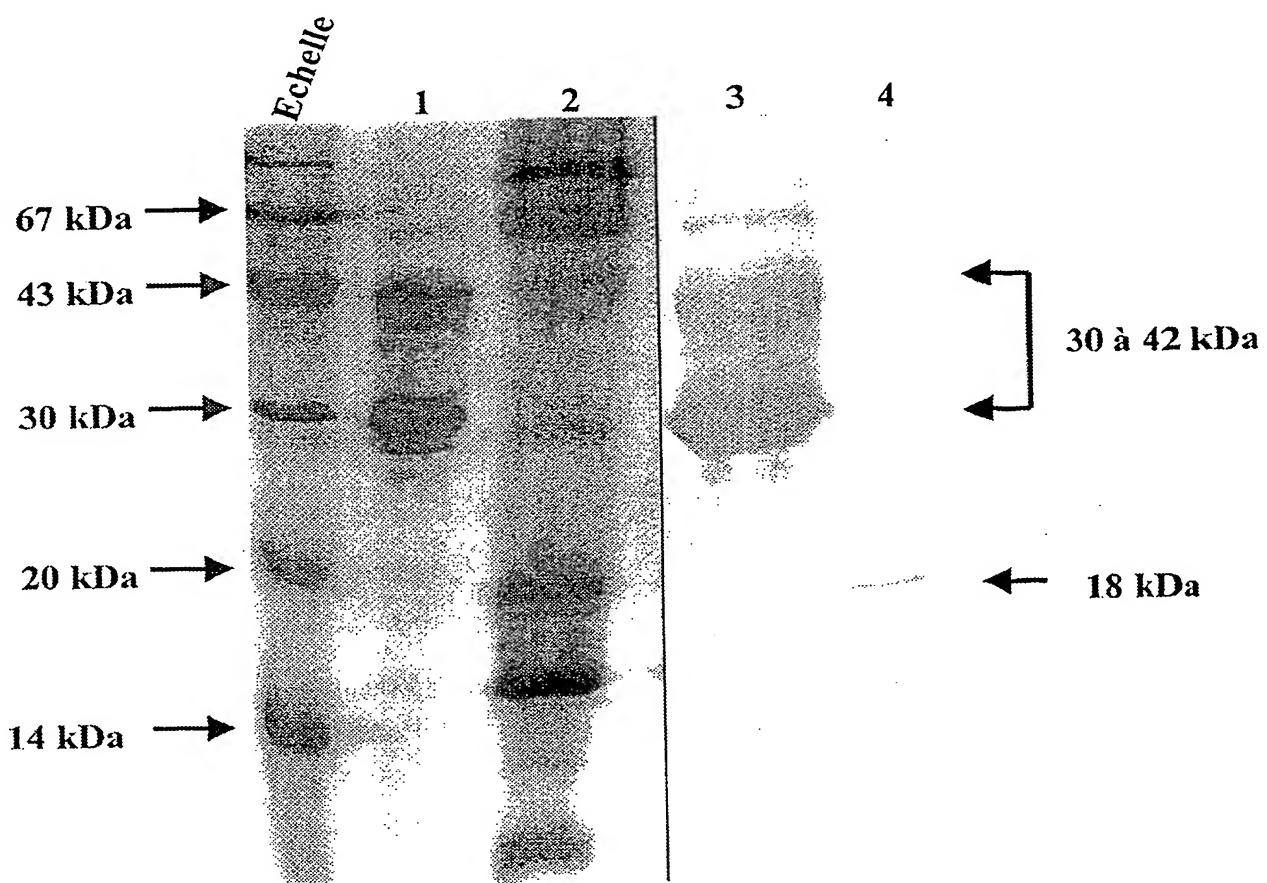


FIG.11

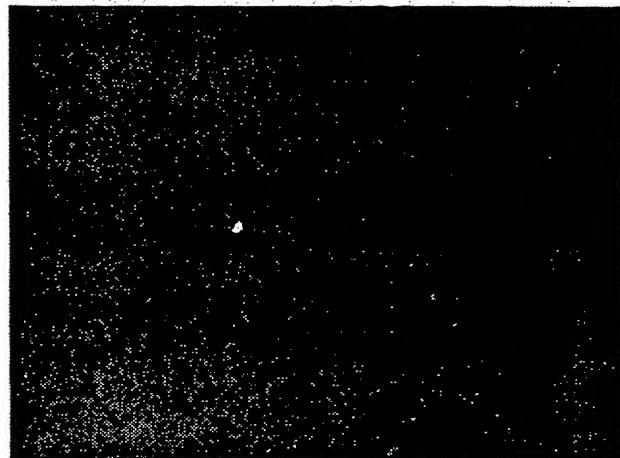
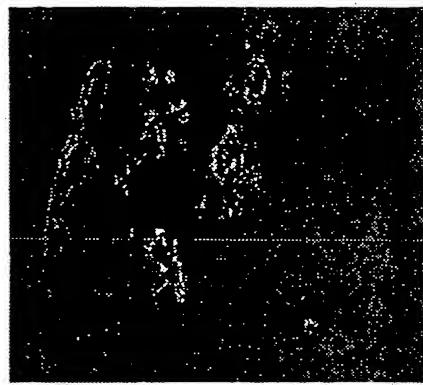
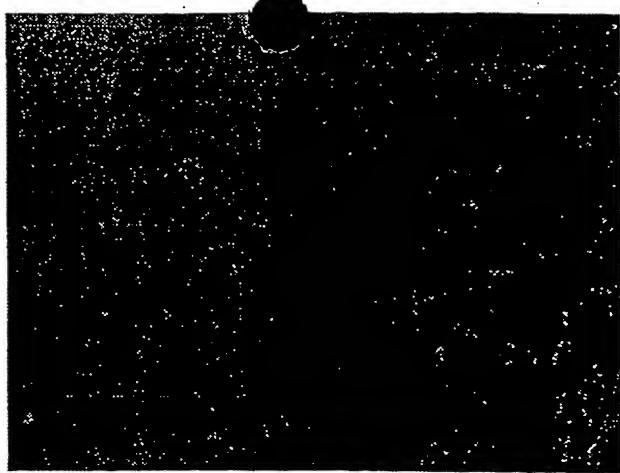
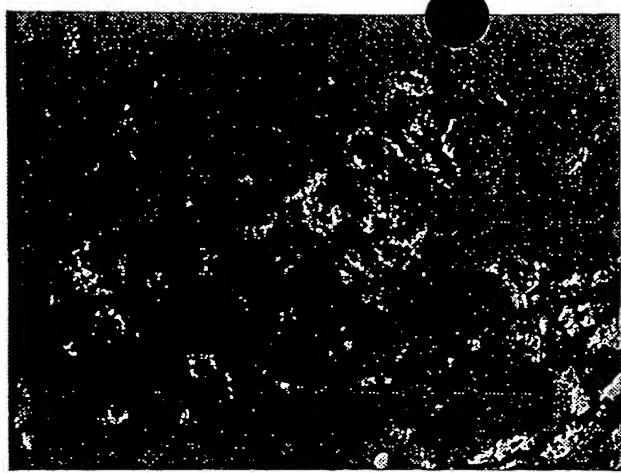


FIG. 12

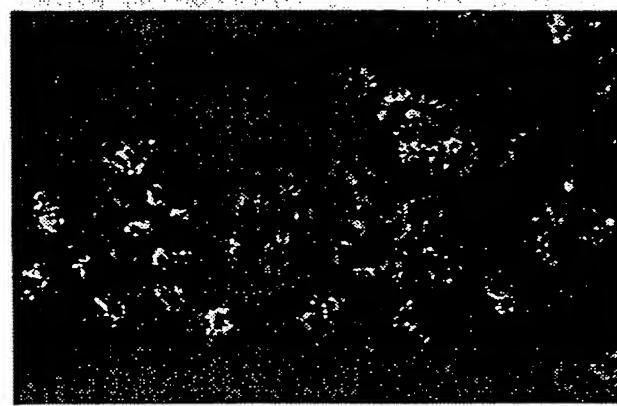
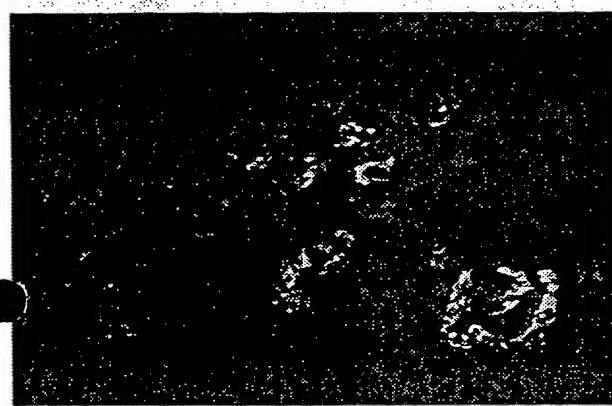
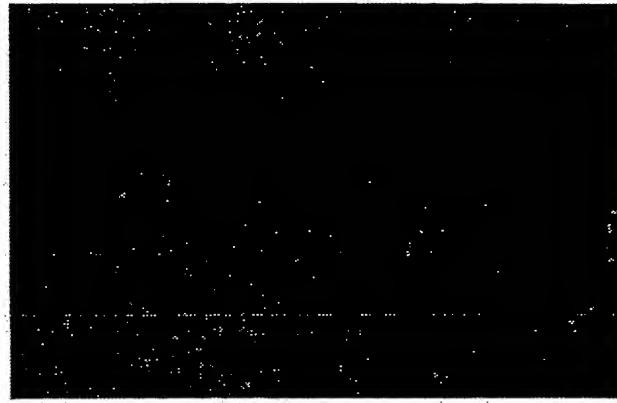
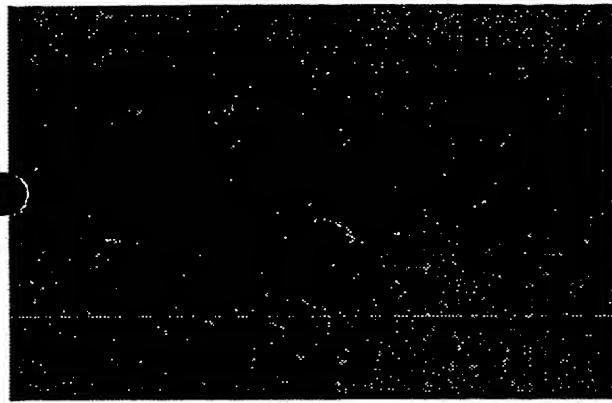
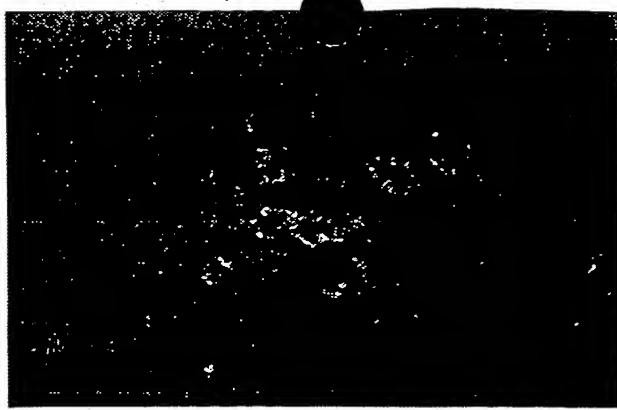
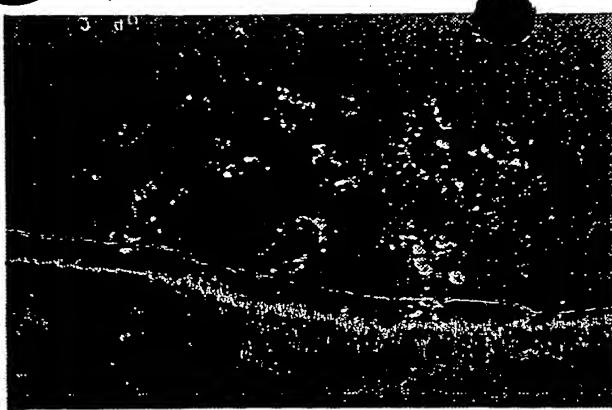


FIG. 13

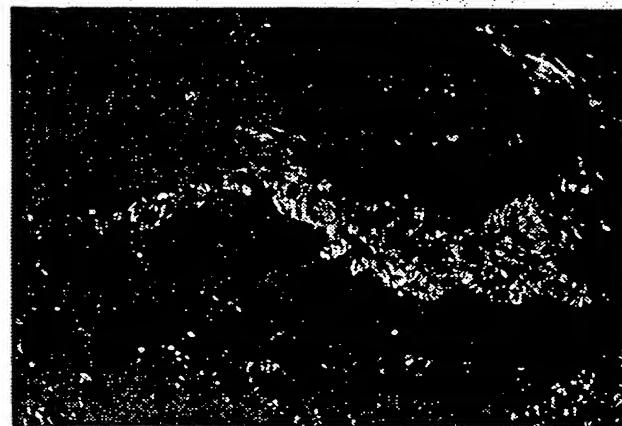
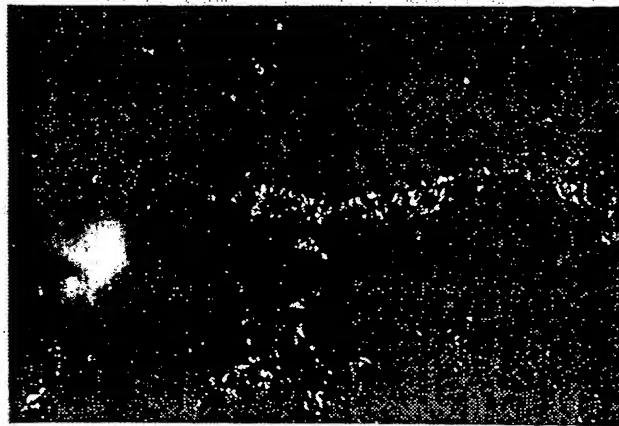
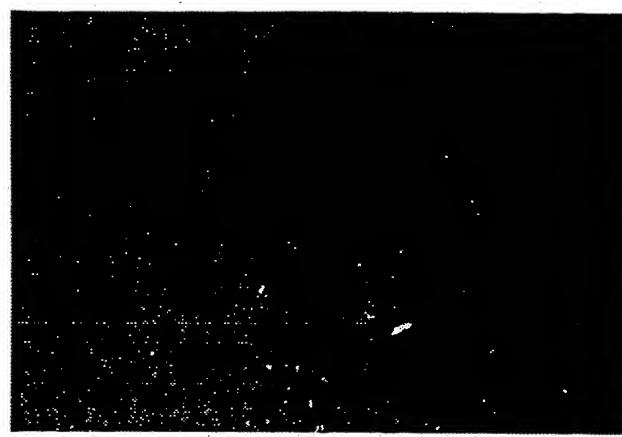
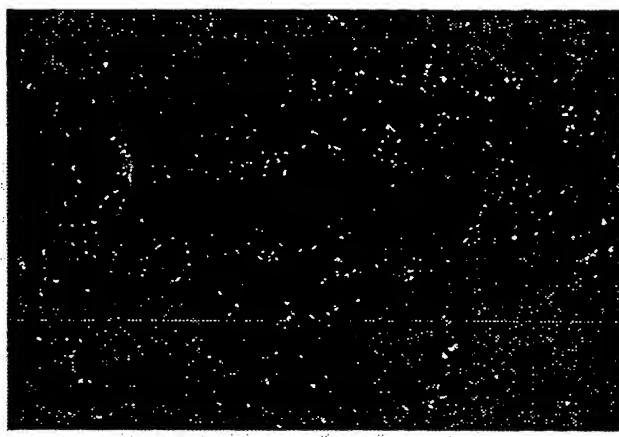
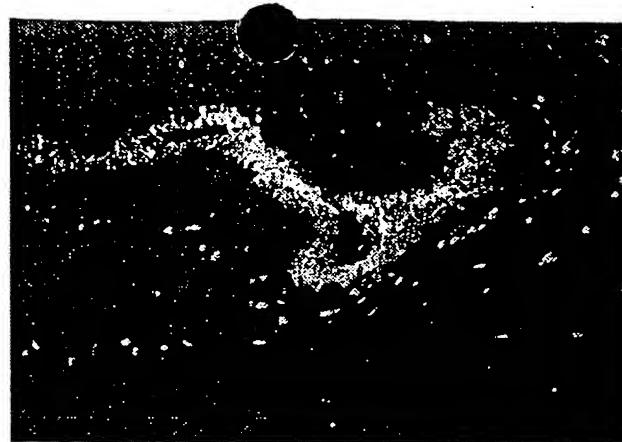
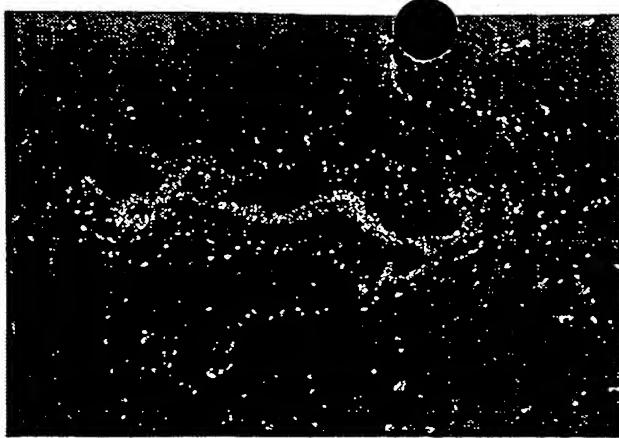


FIG. 14

CABINET MECHANICAL
ORIGINAL

12/14

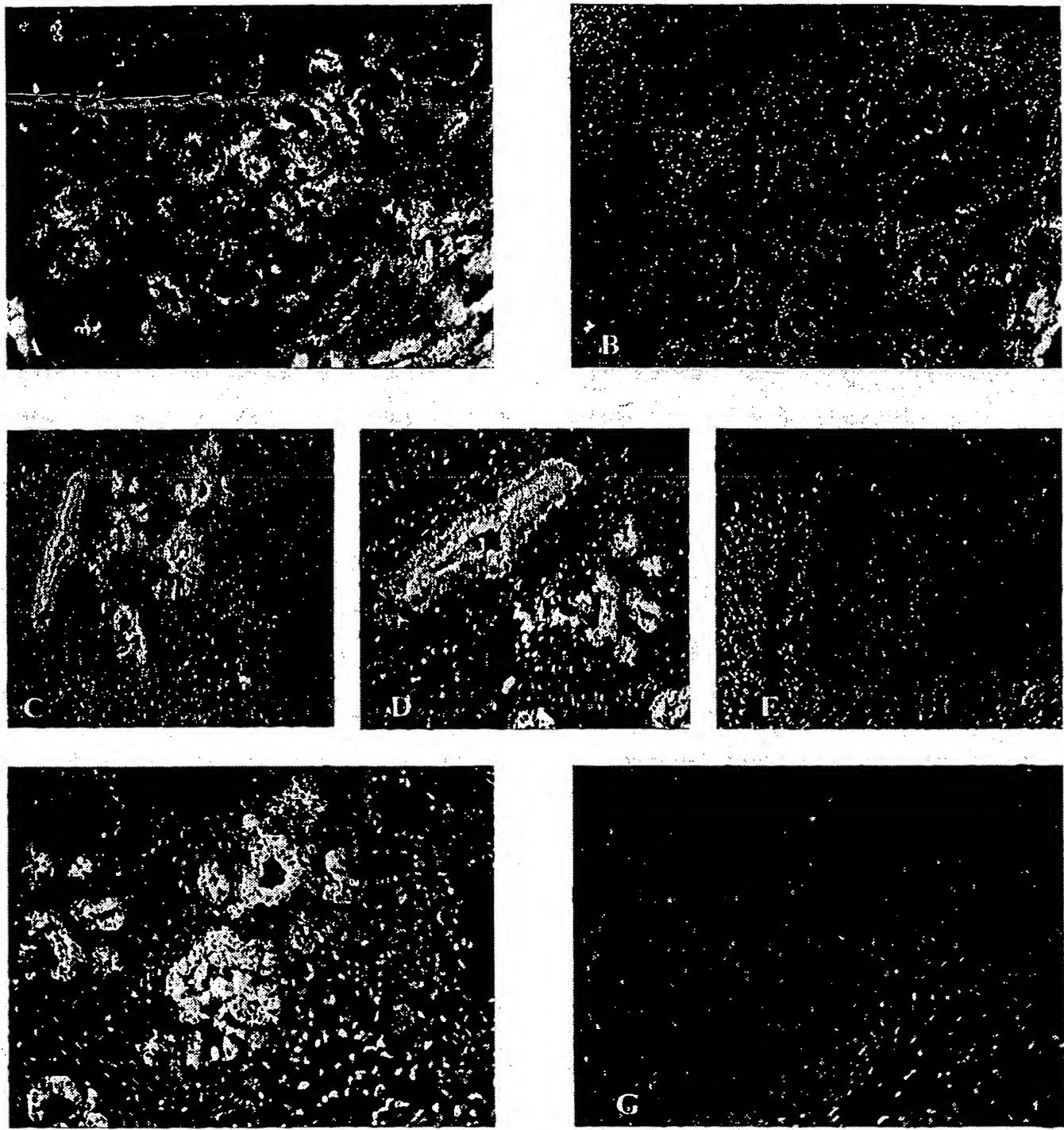


FIGURE 12

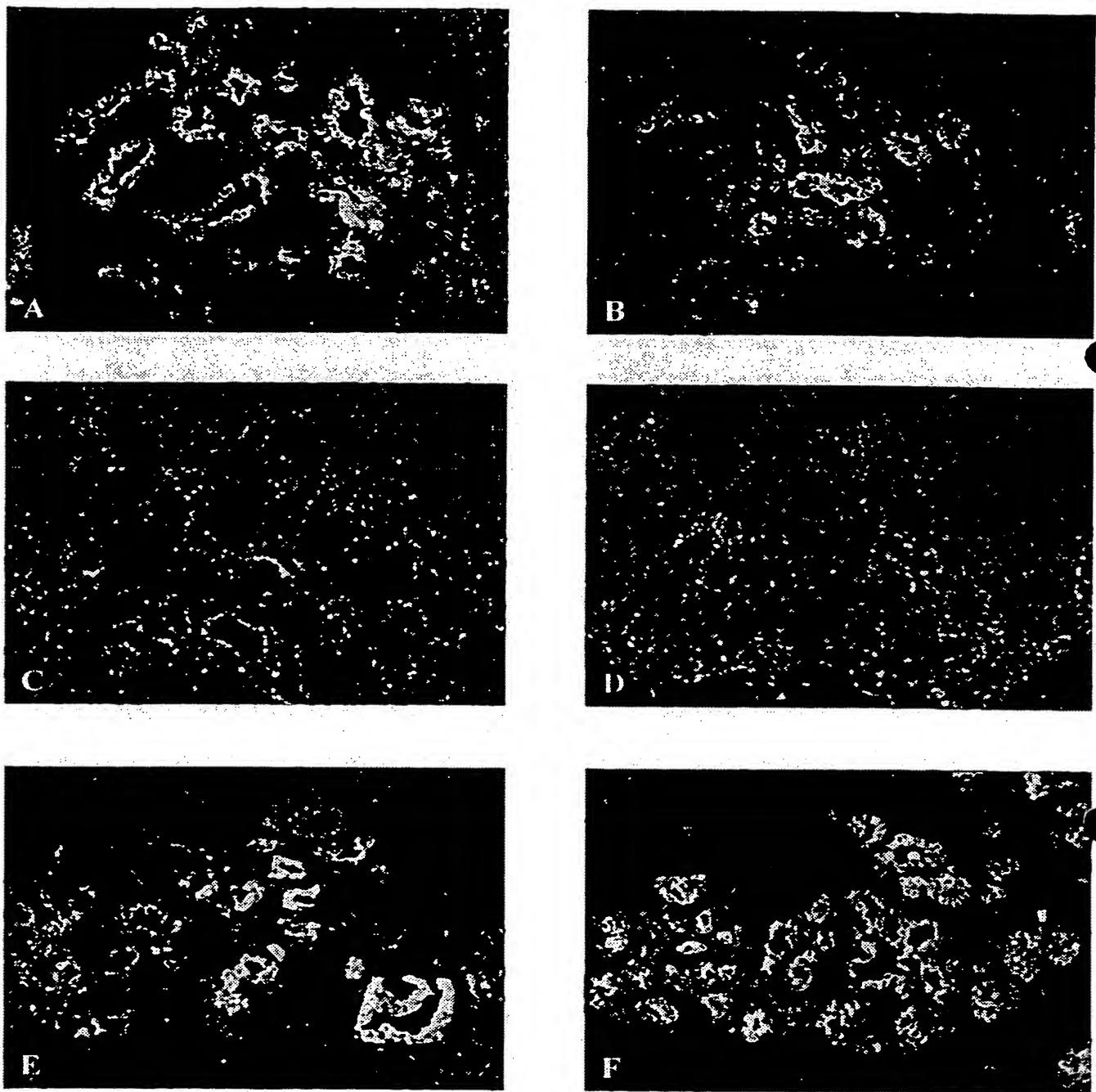


FIGURE 13

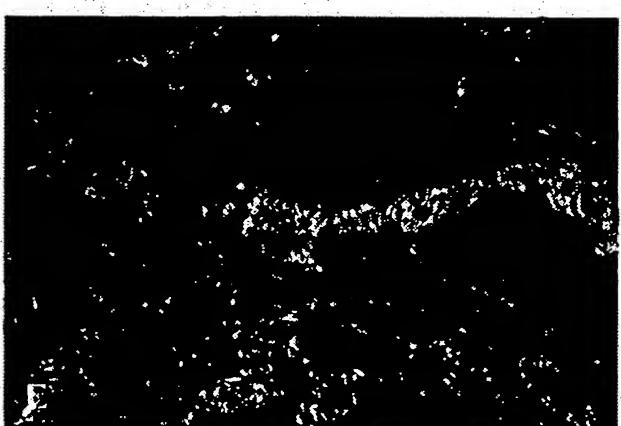
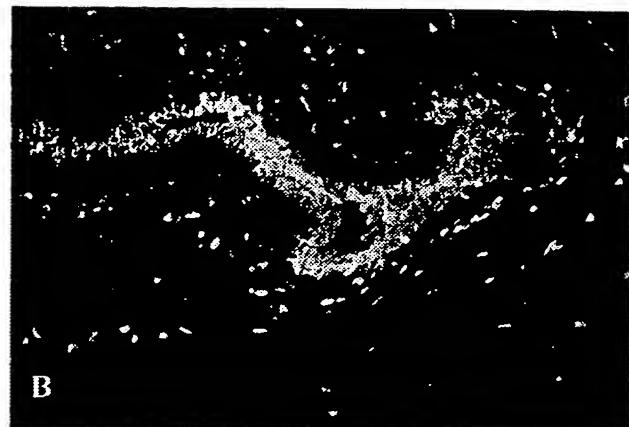
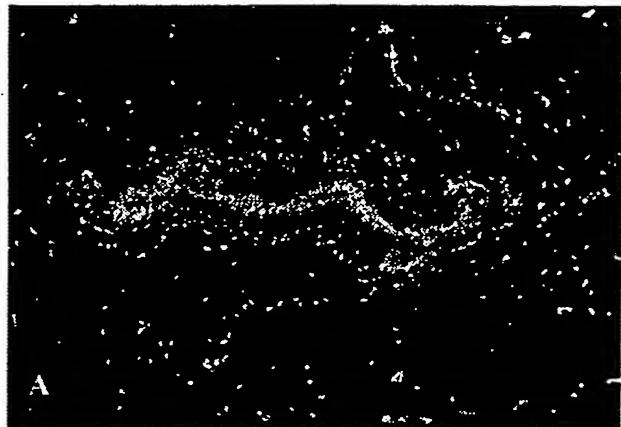


FIGURE 14

